EPA/620/R-94/005
March 1994

Statistical Summary
EMAP-Estuaries
Virginian Province - 1991

by

Steven C. Schimmel
Brian D. Melzian
U.S. Environmental Protection Agency
Environmental Research Laboratory
Narragansett, RI 02882

Daniel E. Campbell
University of Rhode Island
Graduate School of Oceanography

Charles J. Strobel
Sandra J. Benyi
Science Applications International Corporation

Jeffrey S. Rosen
Henry W. Buffum
Computer Sciences Corporation

Virginian Province Manager
Norman I. Rubinstein

Project Officer
Brian Melzian

United States Environmental Protection Agency
Environmental Research Laboratory
27 Tarzwell Drive
Narragansett, RI 02882

APPENDIX A

SAMPLING DESIGN, ECOLOGICAL INDICATORS, AND METHODS

A.1 Region and Estuarine Classification

EMAP-E monitoring is conducted on regional and national scales. Standardized methods are employed, and the entire Virginian Province is sampled simultaneously within a defined "index" time period (July 1 - September 30) to ensure comparability of data within and among sampling years. EMAP-E identified boundaries for 12 estuarine regions (Holland, 1990) based on biogeographic provinces defined previously by NOAA and the U.S. Fish and Wildlife Service using major climatic zones and prevailing major ocean currents (Terrell, 1979) (Figure A-1). The 1990 Virginian Province Demonstration Project and the 1991 survey included the estuarine resources located along the irregular coastline of the mid-Atlantic coast between Cape Cod, MA and Cape Henry, VA, including but not limited to: Buzzards Bay, Narragansett Bay, Long Island Sound, New York/New Jersey Harbors, Delaware Bay, and Chesapeake Bay. Five major rivers within the Province were monitored: the Hudson, the Delaware, the Rappahannock, the Potomac, and the James.

A review of the literature identified potential classification variables that reduced within-class variability. These variables included physical attributes (salinity, sediment type, depth), and extent of pollutant loadings. The use of salinity, sediment type, and pollutant loadings as classification variables *(i.e., a priori* strata) would result in the definition of classes for which areal extents could vary dramatically from year-to-year or even over the index sampling period of EMAP-E. This stratification process requires establishment of a sampling frame prior to sampling; thus misclassification of sample sites within a class should be minimized. Stratification by sediment type, depth, or salinity was considered to be difficult because detailed maps of sediment and water column characteristics were not available or are often unreliable for much of the Virginian Province. These attributes were not used.

A simple classification scheme based on the physical dimensions of an estuary was used to develop three classes -- large estuaries, large tidal rivers, and small estuaries/small tidal rivers. Large estuaries in the Virginian Province were defined as those estuaries greater than 260 km^2 in surface area and with aspect ratios (*i.e.*, length/average width) of less than 18. Large tidal rivers were defined as that portion of the river that is tidally influenced (*i.e.*, detectable tide > 2.5 cm), greater than 260 km^2, and with an aspect ratio of greater than 18. Small estuaries and small tidal rivers were designated as those systems whose surface areas fell between 2.6 km^2 and 260 km^2. These designations excluded estuarine water bodies less than 2.6 km^2 in surface area. These resources were included in the sampling frame by making them a part of the class occupied by their adjacent water body, but were not sampled separately.

Application of the classification scheme based upon geometric dimensions (criteria unlikely to change in reasonable time frames) to the Virginian Province estuarine resources resulted in the identification of 12 large estuaries; 5 large tidal rivers; and 144 small estuaries / small tidal rivers.

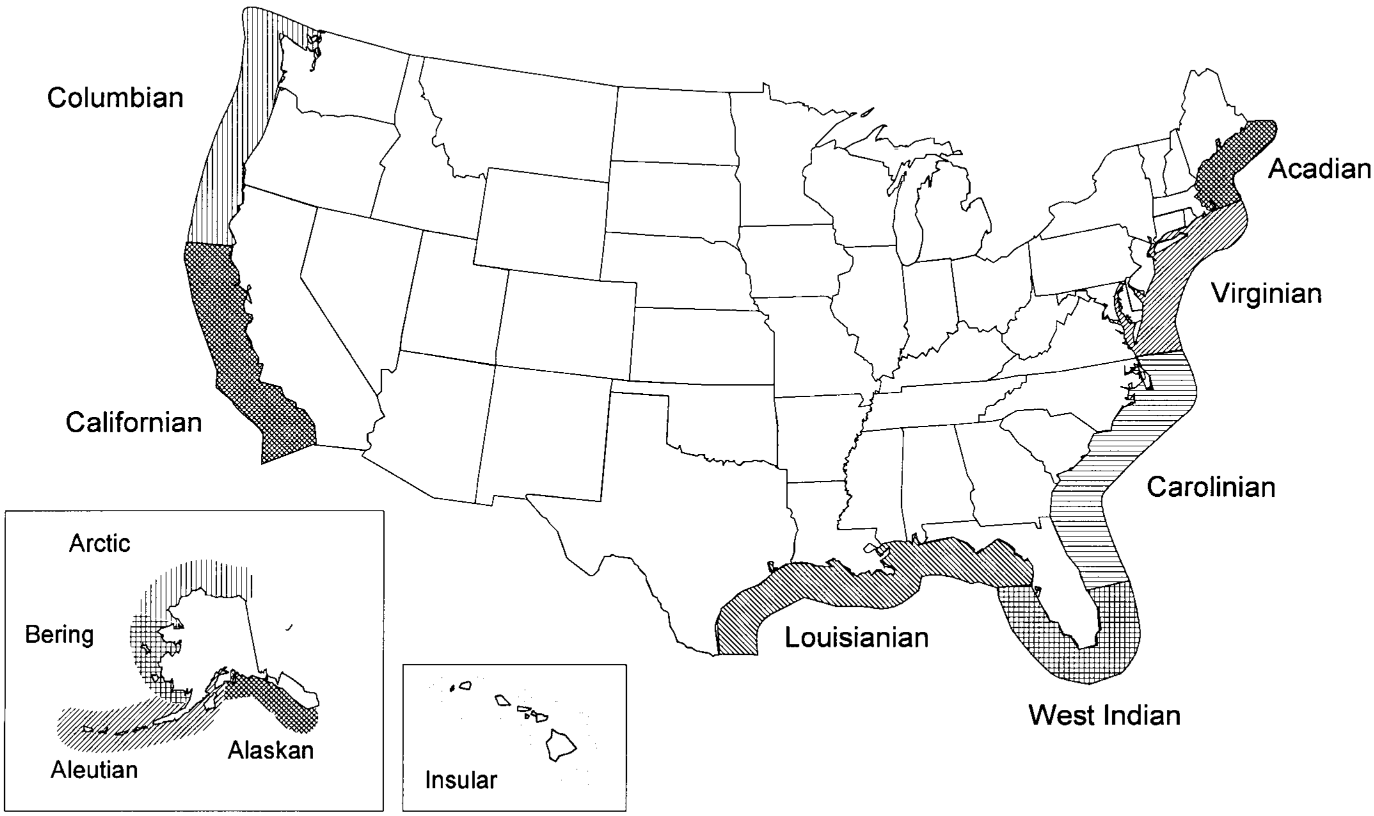


Figure A-1. EMAP-Estuaries biogeographical provinces.

A.2 Sampling Design

Sample collection in the Virginian Province focused on ecological indicators (see Section A.6) during the index sampling period (July 1 - September 30), the period when many estuarine responses to anthropogenic and natural stresses are anticipated to be most severe. The proposed sampling design combines the strengths of systematic and random sampling with an understanding of estuarine ecosystems in order to provide a probability-based estimate of estuarine status in the Virginian Province. In addition, some special-study sites were sampled to collect information for specific hypothesis testing and other specific study objectives. This resulted in sampling five types of sampling sites (stations) for the Virginian Province survey.

- Base Sampling Sites (BSS) are the probability-based sites which form the core of the EMAP-E monitoring design for all provinces, including the Virginian Province. Data collected from these sites are the basis of this preliminary status assessment. There were 102 BSS to be sampled during the 1991 index period. This represents approximately ¼ of the total number of base sites that will be sampled over the four-year cycle.

- Index Sites (IND) were a continuation of a special study initiated in 1990. They are associated with the base sampling sites in small estuarine systems and are located in depositional environments where there is a high probability of sediment contamination or low dissolved oxygen conditions. A total of 29 IND sites were to be monitored in 1991.

- Long-term Trend Sites (LTT) were a select number of 1990 BSS that were revisited in 1991. They will be sampled each year to investigate the within-station annual variability. Eleven (11) LTT sites were monitored in 1991.

- Long-term Trends Spatial Transect (LTS) sites were located along a transect originating at an LTT station. Twelve (12) LTS sites were associated with four transects in 1991 and were all located in the Chesapeake Bay. Three LTS sites were placed along each transect at 0.25, 0.5, and 1.0 statute miles from the associated LTT station to evaluate the spatial variability within a sampling cell.

- Indicator Testing and Evaluation (ITE) sites were initiated in the 1990 Demonstration Project to determine the reliability, sensitivity, and replicability of indicator responses for discriminating between sites with "known" environmental conditions. These sites were selected on the basis of historical information concerning dissolved oxygen concentration and sediment contamination. These sites were used to develop indices and test the discriminatory power of specific indicators. In 1991, six ITE stations were chosen representing our preliminary estimate of degraded, non-degraded or intermediate conditions in the Province. Five of these six stations were filled using existing 1991 BSS sites; one was added as a non-random site and was not used in this assessment. These sites are used to evaluate the performance of "research" indicators. In 1991 the sole research indicators were fish histopathology and redox potential discontinuity (RPD).

A.3 Indicators

EMAP monitoring focuses on indicators of biological response to stress, and uses measures of exposure to stress or contamination as a means for interpreting that response. Traditionally, estuarine monitoring has focused on measures of exposure (*e.g.*, concentrations of contaminants in sediments) and attempted to infer ecological impacts based on laboratory bioassays. The advantage of the ecologically-based approach emphasized in EMAP is that it can be applied to situations where multiple stressors exist, and where natural processes cannot be modeled easily. This is certainly the case in estuarine systems, which are subject to an array of anthropogenic inputs and exhibit a great biotic diversity and complex physical, chemical, and biological interactions.

The implementation plan for the Virginian Province (Schimmel, 1990) listed three general indicator categories for the Demonstration Project: core, developmental and research. These same categories were employed in 1991, although two research indicators were not continued (water column toxicity and relative abundance of and tissue contaminants in large bivalves). Table A-1 lists the indicators retained in the 1991 field survey.

Table A-1. Ecological indicators used in the 1991 Virginian Province Survey.

Category	Indicator
Core	Benthic Species Composition & Biomass Habitat Indicators (Salinity, pH, Temperature,Water Depth, % Silt-Clay)
Developmental	Sediment Contaminants Sediment Toxicity Dissolved Oxygen Concentration Gross Pathology of Fish Contaminants is Fish Tissue Marine Debris Fish Community Composition & Lengths Water Clarity
Research	Histopathology of Fish Apparent RPD

A.4 Field Planning

The success of this complex survey depended upon detailed and complete initial planning. Since 1991 was the second year of sampling in the Virginian Province, many of the documents generated for the 1990 Demonstration Project (with mostly minor modification) served as the basis for protocols, Standard Operating Procedures (SOPs), and various planning documents. With minor exceptions, most of the capital equipment purchased for 1990 was used in 1991. Most contract mechanisms used for field and laboratory analysis activities in 1990 were again used in the succeeding year.

A.4.1 Reconnaissance

An important aspect of planning for the Virginian Province survey was the conduct of desk and field reconnaissance of the Province prior to execution of the survey. The purpose of the reconnaissance was to acquire information that would facilitate the development of logistics plans for field implementation.

The first phase of reconnaissance consisted of plotting all of the 169 stations selected for the Virginian Province by EMAP-E on nautical charts. During this desk exercise, six (6) stations were dropped from the list since they were either in a large estuary class and found to be located on land, or were in a small system that was too shallow throughout to sample. Fifteen (15) sites were located in small estuaries in water less than 2 meters in depth (*i.e.*, the depth needed to deploy Hydrolab meters) and were relocated within the same small estuary. Based on these desk reconnaissance results, a logistics plan (Strobel and Schimmel, 1991a) was developed to sample the remaining 163 sites over a six-week period during July-September 1991, pending the field reconnaissance results.

After the 1991 stations were chosen, an EMAP-E design meeting was held in Columbia, MD to review some of the 1990 results. This meeting produced several changes in the 1991 program. These changes resulted in an overall decrease in station number from 163 to 162.

During field reconnaissance, the crews checked the coordinates, water depth, and access to sites identified as questionable during desk reconnaissance. Seven sites could not be sampled in the field due to insufficient depth. This reduction of seven stations left a total of 155 stations to be sampled in the 1991 field collection effort (the "expected" stations). The logistics plans were modified to reflect these reductions.

In addition to the confirmation of acceptable station conditions, evaluations were made of appropriate boat ramp facilities, motel and restaurant accommodations, Federal Express offices, and dry ice suppliers. Most of the crew chiefs participating in the 1991 survey also participated in the 1990 effort and were, therefore, familiar with many of the sites and the availability of many of these facilities. Therefore, only limited field reconnaissance was required.

A.4.2 Reference Documents

Before training and sampling could be initiated, several reference documents were compiled to guide training, ensure data quality, standardize field and laboratory methods, and provide logistical support to the field teams. These documents include:

- Field Operations and Safety Manual (Strobel and Schimmel, 1991b),
- Quality Assurance Project Plan for the Virginian Province (Valente and Schoenherr, 1991),
- Virginian Province Implementation Plan (Schimmel, 1990),
- Virginian Province Logistics Plan - (Strobel and Schimmel, 1991a),
- EMAP-NC Laboratory Methods Manual (U.S. EPA, 1991),
- 1991 Virginian Province Field Readiness Report (Strobel, 1991).

A.5 Training

Formal training was held at the Environmental Research Laboratory, Narragansett, RI and at the University of Rhode Island Graduate School of Oceanography from May 20 to June 19, 1991. The training was separated into three portions -- Crew Chief Training (late May), Crew Training (June 17-July 3), and Field Certification (July 8-17). A total of 28 participants were trained to conduct sampling in accordance with EMAP protocols. These protocols included vessel navigation, site location, indicator sampling, sample processing, sample shipment, quality control, and logistics. Verification of the crews' understanding of the field protocols was determined by the use of final field certification exercises. These exercises consisted of "typical EMAP-VP" 2-day scenarios comprised of site location, sample collection, sample jar coding, data sheet completion, sample processing and shipping, and electronic data entry and transfer. The crews were required to conduct all components of the sampling activities at each station and were evaluated by senior EMAP-E personnel on their abilities to perform over 100 field and laboratory functions.

All crews were successfully certified. In addition to detailed training on collection methods, shipping procedures, etc., a comprehensive series of first aid and general safety at sea courses were provided to all crew members. These included: cardio-pulmonary resuscitation, general first aid, trauma treatment, use of survival suits, fire extinguishing procedures, etc. Proof of swimming skills was also required of each crew member.

A.6 Indicator Sampling Methods

The EMAP indicator strategy involves four types of ecological indicators (Hunsaker and Carpenter, 1990; Knapp *et al.* 1990): Biotic condition, abiotic condition, habitat, and stressor. Biotic condition indicators are ecological characteristics that integrate the responses of living resources to specific or multiple pollutants and other stresses, and are used by EMAP to assess overall estuarine condition. Abiotic condition indicators quantify pollutant exposure and habitat degradation and are used mainly to identify associations between stresses on the environment and degradation in biotic condition indicators. Habitat indicators provide basic information about the natural environmental gradients. Stressor indicators are used to quantify pollution inputs or stresses and identify the probable sources of pollution exposure. Tables A-2 and A-3 list individual indicators.

Table A-2. Ecological indicators categorized as biotic condition, abiotic condition, and habitat indicators.

Indicator Type	Indicator
Biotic Condition	Benthic Community Composition Benthic Abundance Benthic Biomass Fish Community Composition Fish Lengths Pathology in Fish
Abiotic Condition	Sediment Contaminants Sediment Toxicity Dissolved Oxygen Concentrations Contaminants in Fish Tissue Marine Debris
Habitat	Water Clarity RPD Depth Salinity Temperature Percent Silt-Clay pH Water Depth

Table A-3. Subcomponents of ecological indicators.

Primary Indicator	Subcomponents
Benthos	Total abundance Species composition Species diversity Abundance by species Percentage by taxonomic group Biomass Biomass by taxonomic group
Fish	Total abundance Species composition Species diversity Abundance by species Percentage by taxonomic group Mean length by species
Gross Pathology	Type of disorder
Dissolved Oxygen	Instantaneous at sampling Continuous for 24-hr (15-min intervals)
Sediment Toxicity	*Ampelisca abdita* 10-day test
Sediment Contaminants	23 polycyclic aromatic hydrocarbons 15 metals 15 pesticides 18 PCB congeners Butyltins
Sediment Characters	Percent silt-clay Acid Volatile Sulfides (AVS) Total organic carbon (TOC)
Tissue Contaminants	13 metals 16 pesticides 20 PCB congeners

Descriptions of the methods used for individual indicators have been taken from the Near Coastal Program Plan (Holland, 1990), the Virginian Province Implementation Plan (Schimmel, 1990), the Virginian Province Field Operations and Safety Manual (Strobel and Schimmel, 1991b), and the Near Coastal Laboratory Methods Manual (U.S. EPA, 1991).

A.6.1 Biotic Condition Indicators

A.6.1.1 Benthos

Benthic invertebrate assemblages are composed of diverse taxa with a variety of reproductive modes, feeding guilds, life history characteristics, and physiological tolerances to environmental conditions (Warwick, 1980; Bilyard, 1987). As a result, benthic populations respond to changes in conditions, both natural and anthropogenic, in a variety of ways (Pearson and Rosenberg, 1978; Rhoads *et al.*, 1978; Boesch and Rosenberg, 1981). Responses of some benthic organisms indicate changes in water quality while others indicate changes in sediment quality. Most benthic organisms have limited mobility. They are not as able to avoid exposure to pollution stress as many other estuarine organisms (*e.g.*, fish). Benthic communities have proven to be a reasonable and effective indicator of the extent and magnitude of pollution impacts in estuarine environments (Bilyard, 1987; Holland, *et al.* 1988 and 1989).

Benthic samples for evaluation of species composition, abundance, and biomass were collected at all sampling sites. Samples were collected with a Young-modified van Veen grab that samples a surface area of 440 cm^2. Three (3) grabs were collected at each base, index, or long-term site. A small core (60 cc) was taken from each grab for sediment characterization. The remaining sample was sieved through a 0.5 mm screen using a backwash technique that minimized damage to soft-bodied animals. Samples were preserved in 10% formalin-rose bengal solution and stored for at least 30 days prior to processing to assure proper fixation.

In the laboratory, macrobenthos were transferred from formalin to an ethanol solution and sorted, identified to lowest practical taxonomic level, and counted. Biomass was measured for key taxa and all other taxa were grouped according to taxonomic type (*e.g.*, polychaetes, amphipods, decapods). Shell-free dry weight was determined using an analytical balance with an accuracy of 0.1 mg after drying at 60°C. Large bivalves were shucked prior to determining biomass. Smaller shells were removed by acidification using a 10% HCl solution.

A.6.1.2 Fish

There are several advantages to using fish as a potential indicator of estuarine condition. Because of their dominant position at the upper end of the estuarine food web, fish responses integrate many short-term and small-scale environmental perturbations. Fish are known to respond to most environmental problems of concern in estuaries, including eutrophication, habitat modification, and pathogenic or toxic contamination.

Fish were collected by trawling with a 15 m, high-rise otter trawl with a 2.5-cm mesh cod end. The net was towed for 10 minutes against the tide (if significant tidal current existed) between 0.5 and 1.5 m/s (1-3 knots). All fish caught in the trawl were identified to species and counted; up to 30 fish of a species from each collection were measured to the nearest millimeter.

A maximum of five (5) individuals per target species were retained from each base station trawl for tissue analysis. The specimens were labeled, frozen on dry ice, packaged and shipped to the laboratory where they were stored frozen for subsequent tissue contaminant analysis. In the laboratory, muscle tissue samples were composited by species (by station) and analyzed for the compounds listed in Table A-4.

Individual target species collected in standard trawls were inspected for gross external pathological disorders at all stations where fish were collected. This inspection included checking body surface and fins for skin discoloration, raised scales, white or black spots, ulcers, fin erosion, lumps or growths and condition of the eyes. Specimens with observed gross pathologies were preserved in Dietrich's solution for subsequent laboratory verification and histological examination. At indicator testing sites, all specimens exhibiting gross pathologies, and up to 25 pathology-free specimens of each target species, were preserved for quality control checks of field observations. These fish also received histopathological examinations related to liver lesions, spleen macrophage aggregates, and gill or kidney disfunction (research indicator).

A.6.2 Abiotic Condition Indicators

A.6.2.1 Sediment Collection Procedures

Sediments were collected using the same Young-modified van Veen grab used for benthic invertebrate sampling. The top 2 cm of 6-10 grabs were placed in a teflon mixing bowl and homogenized. Care was taken to avoid collecting sediment adjacent to the edges of the collection device. After approximately 2,000 cc of sediment were collected and completely homogenized, the sediment was distributed among containers for sediment characterization, sediment chemistry, and sediment toxicity testing.

Table A-4. 1991 EMAP Virginian Province: Fish Tissue Chemistry Analytes

Analyte Code	Definition
AG	Silver Concentration in ug/g Dry Weight
AL	Aluminum Concentration in ug/g Dry Weight
AS	Arsenic Concentration in ug/g Dry Weight
CD	Cadmium Concentration in ug/g Dry Weight
CR	Chromium Concentration in ug/g Dry Weight
CU	Copper Concentration in ug/g Dry Weight
FE	Iron Concentration in ug/g Dry Weight
HG	Mercury Concentration in ug/g Dry Weight
MN	Manganese Concentration in ug/g Dry Weight
NI	Nickel Concentration in ug/g Dry Weight
PB	Lead Concentration in ug/g Dry Weight
SE	Selenium Concentration in ug/g Dry Weight
SN	Tin Concentration in ug/g Dry Weight
ZN	Zinc Concentration in ug/g Dry Weight
PCB8	2,4'-dichlorobiphenyl in ng/gram
PCB18	2,2',5-trichlorobiphenyl in ng/gram
PCB28	2,4,4'-trichlorobiphenyl in ng/gram
PCB44	2,2',3,5'-tetrachlorobiphenyl in ng/gram
PCB52	2,2',5,5'-tetrachlorobiphenyl in ng/gram
PCB66	2,3',4,4'-tetrachlorobiphenyl in ng/gram
PCB101	3,3',4,4',5-pentachlorobiphenyl in ng/gram
PCB105	2,2',4,4',5-pentachlorobiphenyl in ng/gram
PCB110/77	2,2',4,5,5'-pentachlorobiphenyl+ 3,3',4,4'-tetrachlorobiphenyl in ng/gram
PCB118	2,3,3',4,4'-pentachlorobiphenyl in ng/gram
PCB126	2,3',4,4',5-pentachlorobiphenyl in ng/gram
PCB128	2,2',3,3',4,4'-hexachlorobiphenyl in ng/gram
PCB138	2,2',3,4,4',5'-hexachlorobiphenyl in ng/gram
PCB153	2,2',4,4',5,5'-hexachlorobiphenyl in ng/gram
PCB170	2,2',3,3',4,4',5-heptachlorobiphenyl in ng/gram
PCB180	2,2',3,4,4',5,5'-heptachlorobiphenyl in ng/gram
PCB187	2,2',3,4',5,5',6-heptachlorobiphenyl in ng/gram
PCB195	2,2',3,3',4,4',5,6-octachlorobiphenyl in ng/gram
PCB206	2,2',3,3',4,4',5,5',6-nonachlorobiphenyl in ng/gram
PCB209	Decachlorobiphenyl in ng/gram
OPDDE	2,4'-DDE DDT and metabolites in ng/gram
PPDDE	4,4'-DDE DDT and metabolites in ng/gram
OPDDD	2,4'-DDD DDT and metabolites in ng/gram
PPDDD	4,4'-DDD DDT and metabolites in ng/gram
OPDDT	2,4'-DDT DDT and metabolites in ng/gram
PPDDT	2,4'-DDT DDT and metabolites in ng/gram
ENDRIN	Endrin in ng/gram
ALDRIN	Aldrin in ng/gram
ALPHACHL	Alpha-Chlordane in ng/gram
TNONCHL	Trans-Nonachlor in ng/gram
DIELDRIN	Dieldrin in ng/gram
HEPTACHL	Heptachlor in ng/gram
HEPTAEPO	Heptachlor epoxide in ng/gram
LINDANE	Lindane (gamma-BHC) in ng/gram
MIREX	Mirex in ng/gram

A.6.2.2 Sediment Characterization

The physical characteristics of estuarine sediments (*e.g.*, grain size) and certain chemical aspects of sediments (*e.g.*, acid volatile sulfide [AVS] content, total organic carbon [TOC] content) influence the distribution of benthic fauna and the accumulation of contaminants in sediments (Rhoads, 1974; Plumb, 1981; DiToro *et al.*, 1991). Sediment silt-clay content was determined to help interpret biotic condition indicator data and sediment contaminant concentrations. AVS and TOC were collected not only as interpretive aids but also as potential covariates for toxic contaminant concentrations.

Subsamples from each benthic grab and contaminant/ toxicity homogenate were retained for grain size determination. A subsample for AVS content was removed from the homogenate. Samples were shipped, on ice, to their respective processing laboratory. Samples for the determination of silt/clay content were sieved using a 63μm mesh sieve. Both the aliquot of the filtrate and the fraction retained on the sieve were dried in an oven at 60°C and weighed to calculate the proportion of silt/clay in the sample.

TOC and AVS concentrations were determined for each site. TOC (from the chemistry sample) was determined by drying a minimum of 5 g wet weight of sediment for 48 hours. Weighed subsamples were ground to fine consistency and acidified to remove sources of inorganic carbon (*e.g.*, shell fragments). The acidified sample was ignited in a furnace at approximately 950°C and the carbon dioxide evoked was measured with an infrared gas analyzer. These peaks were converted to total organic carbon.

The concentration of AVS was determined by the measurement of amorphous or moderately crystalline monosulfides. These substances are important in controlling the bioavailability of metals in anoxic sediments. If the molar ratio of metal to AVS exceeds one, then the metal is potentially bioavailable (DiToro *et al.*, 1990).

The collection methods employed in the 1991 Survey permitted the potential release of sulfides when the materials where processed on-board the sampling vessel and in subsequent shipping. The sample was collected from a homogenized composite (*i.e.*, allowing maximal exposure to oxygen) and shipped on ice to the analytical laboratory. As a result, the accuracy of the AVS measurements could be in doubt although the precision may remain reliable as all samples were treated similarly. Modifications to the collection methods have been determined for the 1992 sampling to prevent a recurrence of these problems.

A.6.2.3 Sediment Contaminants

Metals, organic chemicals, and fine-grained sediments entering estuaries from freshwater inflows, point sources of pollution, and various non-point sources including atmospheric deposition, generally are retained within estuaries and accumulate in the sediments (Turekian, 1977; Forstner and Wittmann, 1981; Schubel and Carter, 1984; Nixon *et al.*, 1986; Hinga, 1988). Samples were collected from a homogenate created during sampling by combining the top 2 cm of sediment from 6-10 sediment grabs. The sediment was placed in clean glass jars with teflon liners or polypropylene containers (for organics and metals analyses, respectively), shipped on ice, and stored frozen in the laboratory prior to analysis for organic contaminants. Sediments were analyzed for the NOAA Status and Trends suite of contaminants (Table A-5).

A.6.2.4 Sediment Toxicity

Sediment toxicity testing is the most direct measure available for determining the toxicity of contaminants in sediments to indigenous biota. It improves upon direct measurement of sediment contaminants because many contaminants are tightly bound to sediment particles or are chemically complexed and, therefore, are not biologically available (U.S. EPA, 1989). Sediment toxicity testing, however, cannot be used to replace direct measurement of the concentrations of contaminants in sediment because such measurements are an important part of interpreting the results of toxicity tests.

Toxicity tests were performed of the composite sediment samples from each station. Tests were conducted using the standard 10-day acute test method (Swartz *et al.*, 1985; ASTM 1991) and the tube-dwelling amphipod *Ampelisca abdita*.

Table A-5. 1991 EMAP Virginian Province: Sediment Chemistry Analytes

Analyte Code	Definition
TOC	Total Organic Carbon Concentration in µg/g Dry Weight
AG	Silver Concentration in µg/g Dry Weight
AL	Aluminum Concentration in µg/g Dry Weight
AS	Arsenic Concentration in µg/g Dry Weight
CD	Cadmium Concentration in µg/g Dry Weight
CR	Chromium Concentration in µg/g Dry Weight
CU	Copper Concentration in µg/g Dry Weight
FE	Iron Concentration in µg/g Dry Weight
HG	Mercury Concentration in µg/g Dry Weight
MN	Manganese Concentration in µg/g Dry Weight
NI	Nickel Concentration in µg/g Dry Weight
PB	Lead Concentration in µg/g Dry Weight
SB	Antimony Concentration in µg/g Dry Weight
SE	Selenium Concentration in µg/g Dry Weight
SN	Tin Concentration in µg/g Dry Weight
ZN	Zinc Concentration in µg/g Dry Weight
PCB8	2,4'-dichlorobiphenyl in ng/gram
PCB18	2,2',5-trichlorobiphenyl in ng/gram
PCB28	2,4,4'-trichlorobiphenyl in ng/gram
PCB44	2,2',3,5'-tetrachlorobiphenyl in ng/gram
PCB52	2,2',5,5'-tetrachlorobiphenyl in ng/gram
PCB66	2,3',4,4'-tetrachlorobiphenyl in ng/gram
PCB101	3,3',4,4',5-pentachlorobiphenyl in ng/gram
PCB105	2,2',4,4',5-pentachlorobiphenyl in ng/gram
PCB118	2,3,3',4,4'-pentachlorobiphenyl in ng/gram
PCB128	2,2',3,3',4,4'-hexachlorobiphenyl in ng/gram
PCB138	2,2',3,4,4',5'-hexachlorobiphenyl in ng/gram
PCB153	2,2',4,4',5,5'-hexachlorobiphenyl in ng/gram
PCB170	2,2',3,3',4,4',5-heptachlorobiphenyl in ng/gram
PCB180	2,2',3,4,4',5,5'-heptachlorobiphenyl in ng/gram
PCB187	2,2',3,4',5,5',6-heptachlorobiphenyl in ng/gram
PCB195	2,2',3,3',4,4',5,6-octachlorobiphenyl in ng/gram
PCB206	2,2',3,3',4,4',5,5',6-nonachlorobiphenyl in ng/gram
PCB209	Decachlorobiphenyl in ng/gram
MBT	Mono-butyl Tin in ng/gram
DBT	Di-butyl Tin in ng/gram
TBT	Tri-butyl Tin in ng/gram
OPDDE	2,4'-DDE DDT and metabolites in ng/gram
PPDDE	4,4'-DDE DDT and metabolites in ng/gram
OPDDD	2,4'-DDD DDT and metabolites in ng/gram
PPDDD	4,4'-DDD DDT and metabolites in ng/gram
OPDDT	2,4'-DDT DDT and metabolites in ng/gram
PPDDT	2,4'-DDT DDT and metabolites in ng/gram

(Continued)

Table A-5 (continued).

Analyte Code	Definition
ALDRIN	Aldrin in ng/gram
ALPHACHL	Alpha-Chlordane in ng/gram
TNONCHL	Trans-Nonachlor in ng/gram
DIELDRIN	Dieldrin in ng/gram
HEPTACHL	Heptachlor in ng/gram
HEPTAEPO	Heptachlor epoxide in ng/gram
HEXACHL	Hexachlorobenzene in ng/gram
LINDANE	Lindane (gamma-BHC) in ng/gram
MIREX	Mirex in ng/gram
NAPH	Naphthalene in ng/gram
MENAP2	2-methylnaphthalene in ng/gram
MENAP1	1-methylnaphthalene in ng/gram
BIPHENYL	Biphenyl in ng/gram
DIMETH	2,6-dimethylnaphthalene in ng/gram
ACENTHY	Acenaphthlylene in ng/gram
ACENTHE	Acenaphthene in ng/gram
TRIMETH	2,3,5-trimethylnaphthalene in ng/gram
FLUORENE	Fluorene in ng/gram
PHENANTH	Phenanthrene in ng/gram
ANTHRA	Anthracene in ng/gram
MEPHEN1	1-methylphenanthrene in ng/gram
FLUORANT	Fluoranthene in ng/gram
PYRENE	Pyrene in ng/gram
BENANTH	Benz(a)anthracene in ng/gram
CHRYSENE	Chrysene in ng/gram
BENZOBFL	Benzo(b)fluoranthene in ng/gram
BENZOKFL	Benzo(k)fluoranthene in ng/gram
BENAPY	Benzo(a)pyrene in ng/gram
BENEPY	Benzo(e)pyrene in ng/gram
PERYLENE	Perylene in ng/gram
INDENO	Ideno(1,2,3-c,d)pyrene in ng/gram
DIBENZ	Dibenz(a,h)anthracene in ng/gram
BENZOP	Benzo(g,h,i)perylene in ng/gram
SAND_PC	Sand Content (%)
SICL_PC	Silt-Clay Content (%)

A.6.2.5 Dissolved Oxygen

Dissolved oxygen (DO) is a fundamental requirement for maintenance of populations of benthos, fish, shellfish, and other estuarine biota. DO concentrations are affected by environmental stresses, such as point and non-point discharges of nutrients or oxygen-demanding materials (*e.g.*, particulates, dissolved organic matter). In addition, stresses that occur in conjunction with low DO concentrations may be even more detrimental to biota (*e.g.*, exposure to hydrogen sulfide, decreased resistance to disease and contaminants). DO levels are highly variable over time, fluctuating widely due to tidal action, wind stress, and biological activity (Kemp and Boynton, 1980; Welsh and Eller, 1991).

Dissolved oxygen was sampled in three ways during the 1991 Virginian Province survey: 1) instantaneous water column profiles using a SeaBird model SBE 25 CTD, (2) point in time bottom oxygen conditions with a YSI (model 58) oxygen meter and the SeaBird CTD, and 3) continuous 24-72 hr measurements of bottom concentrations using a Hydrolab DataSonde 3 data logging array. The first two measurements were taken at all sites, and the continuous measurements were taken at base stations (BSS) only.

The Hydrolab DataSonde 3 data logger deployed at each Base site for 24-72 hours collected continuous DO data at 15-min intervals. The DataSonde 3 also collected salinity, temperature, water depth, and pH data. The instruments were calibrated prior to every deployment, and were checked on-board ship immediately prior to deployment by comparison to the YSI oxygen meter, and following retrieval. These instruments were deployed approximately 1 m from the bottom. The stored data were downloaded to a computer and the unit was serviced and recalibrated for subsequent deployment at another site.

A.6.3 Habitat Indicators

Habitat indicators provide basic information about the natural environmental setting. Habitat indicator data included in this report include water depth, salinity, temperature, pH, water clarity, and sediment silt/clay content.

All water quality measurements were made using the Seabird model SBE 25 CTD (described earlier). This unit was equipped with probes to measure salinity, temperature, depth, pH, DO, light transmission, fluorescence, and PAR. All CTD data were downloaded to a computer in the field for review and storage.

Measurements of water clarity are incorporated into the CTD casts that were performed at each station. Included in the CTD instrumentation package are a SeaTech transmissometer and a Biospherical PAR (Photosynthetically Active Radiation) sensor. As the CTD is lowered through the water column, transmissivity and PAR data are continually logged. The main concern in obtaining these measurements is that the lenses on the transmissometer are cleaned before each use.

Surficial water samples were collected at all stations for determination of Total Suspended Solids (TSS). Samples were refrigerated, returned to the laboratory, filtered through a glass-fiber filter, dried and weighed.

Sediment silt/clay content was measured on samples taken from the surficial sediment (top two cm) homogenate from which chemistry and toxicity samples were also removed.

The kinds and amounts of floating and submerged (*i.e.*, collected in trawls) marine debris were noted at all stations. Debris was categorized as paper, plastics, metal, glass, wood, and other wastes. Only debris of anthropogenic origin was included. Wastes that were comprised of composited materials (*e.g.*, metal, wood, and plastic) were categorized based on their dominant component.

A.7 Data Collection and Sample Tracking

Each field crew was supplied with two portable computers and appropriate software to facilitate electronic recording of the data, data transfer, and sample tracking. All samples, shipments, and equipment were labelled with bar-coded labels to facilitate sample tracking and reduce transcription errors. Field computers were equipped with bar code readers to record sample identification numbers. Receiving laboratories were also equipped with bar code readers to facilitate the receiving process and to rapidly convey information concerning lost or damaged shipments.

Copies of all data entered into the field computer were stored on the hard disk and copied to diskettes. Information on the hard disk was transferred daily via commercial carrier phone lines to the Information Management Center at ERL-Narragansett (RI).

Backup diskettes and hard-copy data sheets were shipped weekly to the Center.

All transferred data were examined within 24-48 hours of collection by EMAP-E personnel. Errors were brought to the attention of the field crews for correction and resampling, if required. All electronic data were checked against paper data forms for verification. Further information on the details of the Near Coastal data management systems are presented in Rosen *et al.* (1990).

A.8 Analytical Methods For Statistical Summary

Three types of analyses were conducted for this report: 1) direct descriptions of measured indicators, 2) development of modified or adjusted indicators (*e.g.*, metal contaminants in sediments), and 3) development of indices based on directly measured indicators. These analyses are documented in a Virginian Province 1990 Demonstration Report (Weisberg *et al.*, 1993).

A.8.1 Cumulative Distribution Functions (CDFs)

All ecological indicators collected during the 1991 Virginian Province survey were characterized using Cumulative Distribution Functions (CDFs). These functions describe the full distribution of indicators in relation to their areal extent within the Province. All observations are weighted based upon surface area associated with each sampling site. The area associated with each sampling unit in large estuaries was equal to the hexagonal spaces created by the EMAP grid (70 km^2). For tidal river and small estuary classes, the area associated with each sampling segment was determined using the ARC/INFO data model which produces areal and perimeter estimates. For the tidal river class ARC/INFO version 5.0 (ESRI) was used to delineate the extent of each segment on a 1:100,000 digital line graph. The total areas associated with the three classes is: large estuaries - 16,097 km^2; large tidal rivers - 2,602 km^2; and small estuaries - 4,875 km^2.

To generate estimates across classes (strata), weights for stations within each class were adjusted so that the total of the weights for that class was equal to the total area represented by the stations within that class. This adjustment primarily affects the small estuary class because this resource is sub-sampled from the list of all possible systems. Only 29 of the 144 small estuarine class systems were sampled in 1991 representing 18% of the total surface area of the small estuary/ small tidal river class.

A.8.2 Adjustment To Known Covariates

In several cases, variability in observed indicators might reflect relationships to known habitat or control variables. Examples of these relationships are: variation in estuarine biota resulting from sampling throughout the salinity gradient; variation in sediment toxicity tests with different mortalities associated with the controls; and variation in sediment metals observed at a site resulting from variations in the amount of natural crustal materials at the site. In all these cases, the observed data must be adjusted in order to construct CDFs or to compare observations from different locations.

A.8.2.1 Adjustment for Natural Habitat Gradients

Estuarine biota are largely controlled by their environmental settings, both natural and anthropogenic. Natural gradients, particularly in salinity and silt-clay content, are common in estuaries. Many estuarine organisms may represent overlapping discrete distributions along these gradients. Thus, normalization of ecological measures over habitat gradients is a common tool used to interpret information when such normalization is necessary.

Many ecological variables are significantly correlated with natural gradients (*i.e.*, salinity, silt-clay content, water depth). However, these correlations often explain very little of the total variation observed in that variable. The 1990 EMAP-E effort in the Virginian Province (Weisberg *et. al.*, 1993) suggested that benthic distributions (*e.g.*, number of species, percentage community composition, biomass) needed to be adjusted for salinity gradients before inclusion in the "Benthic Index" (see Appendix B) because these gradients explained greater than 25% of the total variation in these benthic indicators. Similar analyses using the 1990 Virginian Province data showed that while many other benthic and fish indicators were significantly correlated to habitat gradients; none of these correlations accounted for more than 15% of the total variability. As a result, only the expected number of benthic species used in the benthic index was adjusted for salinity.

A.8.2.2 Adjustment for Experimental Controls

Estimates of the area in the Virginian Province containing toxic sediments were based on the results of toxicity tests using the amphipod, *Ampelisca abdita*. For this summary, a relative measure of toxicity was created to facilitate comparisons between sites over a series of bioassays. This adjustment is necessary because control mortalities vary among test series. Sediments were determined to be toxic if: the survival of the test organism in test sediments was less than or equal to 80% of the survival observed in clean, control sediments; if the survivals in test and control sediments were significantly different ($p < 0.05$); and if survival in control sediments was ≥ 85%. This results in an adjustment to the observed survival rates in test sediments that accounts for variability due to differences in the controls for individual bioassays. These criteria are consistent with those established in U.S. EPA/ACE (1991).

A.8.2.3 Adjustment for Natural Crustal Properties

The extent to which anthropogenic activities have affected concentrations of metals in sediments is complicated by the natural variation of concentrations due to differing particle size distributions in sediments. Because of surface adsorptive and complexation processes, fine-grained sediments will naturally have higher trace metal concentrations than coarse sediments. In some studies, *e.g.*, the National Status and Trends program, reported concentrations are adjusted for this variation by normalizing the concentrations by the fine-grained fraction determined separately. As an alterative to actual size-fractionation measurements, a number of authors (Windom *et al.*, 1989; and Schropp *et al.*, 1990) have determined relationships between sediment concentrations of trace metals and other elements indicative of fine-grained crustally-derived material, *e.g.*, aluminum, iron and manganese. The most commonly used of these indicator elements is aluminum, due to its large natural abundance, freedom from common anthropogenic contaminant sources and significant correlation with both the fine-grained fraction and trace metal concentrations in clean, un-impacted sediments. The correlation between aluminum and trace metals in fine-grained sedimentary material has a geochemical basis related to the composition of crustal material from which the fine particles are derived and the natural adsorption and complexation processes occurring during "weathering" of the crustal material. Once background sediment metal-aluminum relationships have been determined, concentrations of metals expected from background material can be subtracted from total metal concentrations, allowing residual, presumably anthropogenic, contributions to be assessed.

Background metal-aluminum relationships are derived by linear regression of sediment concentrations of each element against aluminum concentrations in the same sediment. Some investigators have used log-transformed metals concentrations in the regression analyses. Such transformations do not improve correlation of the metals-aluminum concentrations of this data set. Furthermore, linear regressions provide direct correlation with the physical mixing and geochemical factors noted above which affect the overall concentration of metals in sediments. This correlation is lost when the concentrations are transformed. Consequently, no data transformations were performed prior to regression analysis.

Use of linear regression to determine metal-aluminum relationships in background sedimentary material can only be successful if the sediments do not include contributions from sources other than natural background sediments. The data sets used in this study were statistically screened to eliminate samples which might contain additional source materials. This was accomplished by performing linear regressions of concentrations of aluminum against each metal. The residuals (the differences between the measured concentrations and those predicted from the regression) were then tested for normal distribution. If the residuals were found not to be normally distributed, samples which had studentized residual values greater than 2 were eliminated from the data set . Regression of the reduced data set was repeated and the residuals tested again for normal distribution. This process was repeated for each metal until residuals from the regressions were all normally distributed, at which point the remaining samples were assumed to represent natural, background sediments.

The regression relationships derived for the background sediments were then applied to the original data set. Samples with trace metal concentrations exceeding the upper 95% confidence limit for that metal's regression against aluminum were designated as enriched. It should be noted that no assessment was made as to the magnitude of enriched; metal concentrations might be only slightly above the 95%confidence limit or might exceed the limit by factors of 10-100. The categorization "enriched" was applied to any sediment with a metal concentration higher than that expected from the background sediment aluminum metal relationship at the 95% confidence level.

A.9 Procedures for the Calculation of Confidence Intervals

The approximate 95% confidence intervals for the Province were calculated based on the assumption that the CDF estimates were distributed normally. The confidence intervals were obtained by adding and subtracting 1.96 times the estimated standard error (square root of the variance) to the estimated CDF value.

For small estuarine systems, estimates of CDFs and associated variances were computed based on a random selection of small systems within the Province, with replicate samples taken from a subset of the selected systems (Cochran, 1977). The resulting CDF estimate is:

$$\hat{P}_{Sx} = \frac{\sum_{i=1}^{n} A_i \bar{y}_i}{\sum_{i=1}^{n} A_i}$$

where,

$\hat{P}_{Sx}$ = CDF estimate for value x

$\bar{y}_i = \frac{1}{m_i} \sum_{i=1}^{m_i} y_{ij}$

m_i = number of samples at small system i

A_i = area of small system i

$y_{ij} = \begin{cases} 1 \text{ if response is less than } x \\ 0 \text{ otherwise} \end{cases}$

n = number of small systems sampled

Since replicate samples were only obtained at a subset of the sampled small estuarine systems, the formula for the estimated variance taken from Cochran (1977 eq. 11.30) was modified to produce the following estimate of the approximate mean squared error (MSE) of the CDF estimate:

$$MSE(\hat{P}_{Sx}) = \frac{\frac{N^2}{n}(1-f_1)\frac{\sum_{i=1}^{n} A_i^2(\bar{y}_i-\hat{P}_{sx})^2}{n-1} + \frac{N}{n*}\sum_{i=1}^{n*}\frac{A_i^2 S_{2i}^2}{m_i}}{A^2}$$

where,

$f_1 = n/N$

$n*$ = number small systems with replicate samples

$S_{2i}^2 = \frac{\sum_{i=1}^{m_i} (y_{ij}-\bar{y}_i)^2}{m_i-1}$

A = the total area of small systems in the Province (4,875 km^2)

N = number small systems in Province (144)

Estimates of CDFs for large tidal rivers were obtained by applying Horvitz-Thompson estimation (Cochran, 1977) with selection probabilities being inversely related to station area. Estimates of CDFs were:

$$\hat{P}_{Tx} = \frac{1}{A}\sum_{i=1}^{n}\frac{y_i}{\pi_i}$$

where,

$\hat{P}_{Tx}$ = Estimate CDF at value x

$y_i = \begin{cases} 1 \text{ if response is less than } x \\ 0 \text{ otherwise} \end{cases}$

π_i = inclusion probability for station i (1/area)

A = total area of sampled tidal rivers

n = number of stations sampled

To achieve unbiased estimates of variance, joint event probabilities(π_{ij}) must be non-zero. For large tidal river sampling, many joint event probabilities are zero and other joint event probabilities are unknown; therefore, an approximate variance estimate for the CDF estimates was obtained by applying the Yates-Grundy estimate of variance (Cochran, 1977) and using approximate joint event probabilities (Stevens *et al.*, 1991):

$$v\hat{a}r(\hat{P}_{Tx}) = \frac{1}{A^2}\sum_{i=1}^{n}\sum_{j>i}^{n}\left(\frac{\pi_i\pi_j-\pi_{ij}}{\pi_{ij}}\right)\left(\frac{y_i}{\pi_i}-\frac{y_j}{\pi_j}\right)^2$$

where,

$$\pi_{ij} = \text{probability that sites } i \text{ and } j \text{ are selected for sampling}$$

and

$$\pi_{ij} = \frac{2(n-1)\pi_i\pi_j}{2n-\pi_i-\pi_j}.$$

Estimates of CDFs for large systems ($\hat{P}_{Lx}$) were also obtained by applying Horvitz-Thompson estimation with selection probabilities being inversely related to station area. Areas for all base stations were assumed to be 70 km^2. Formulae for the CDF estimates and corresponding variances are analogous to those presented for large tidal rivers.

Estimates of CDFs for a particular geographic system within the Province (*e.g.*, Chesapeake Bay) were obtained by applying the above procedures to the small estuarine systems, tidal rivers, and large estuaries sampled within that geographic system. Estimates of the CDFs for the entire Province or for a geographic system within the Province were computed as weighted averages of the relevant station class CDFs:

$$\hat{P}_X = W_S\hat{P}_{Sx} + W_T\hat{P}_{Tx} + W_L\hat{P}_{Lx}$$

where,

W_S = Relative area of small systems

W_T = Relative area of tidal rivers

W_L = Relative area of large estuaries

In applying these procedures, variance estimation was based on the assumption of a fixed sample size within each resource class. For large tidal rivers and large estuaries, the sample size is a random element depending on the position of the sampling grid. This variance component has not been incorporated into the estimation of variances of CDFs.

APPENDIX B

CALCULATION OF THE BENTHIC INDEX

B.1 BACKGROUND

Biotic condition indicators are characteristics of the environment that provide quantitative evidence of the status of ecological resources and biological integrity of the site from which they are drawn (Messer, 1990). Ecosystems with a high degree of biotic integrity (*i.e.*, healthy ecosystems) are composed of balanced populations of indigenous organisms with species compositions, diversity, and functional organization comparable to natural habitats (Karr and Dudley, 1981; Karr *et al.*, 1986). Biotic condition indicators could include measurements of the kinds and abundances of biota present, the health of individual organisms, and the sustainability of critical ecological processes. They are the empirical data collected by EMAP-E that are integrated into indices that track the status and trends in ecological integrity.

One category of biotic condition indicators which was measured during the 1991 Virginian Province effort was the benthic indicators. Results are presented for numbers of species, total abundance, and an integrated "Benthic Index".

Benthic invertebrates are the major trophic link between primary producers and higher trophic levels, including fish, shellfish, birds and other wildlife (Carriker, 1967; Rhoads, 1974). They are a particularly important source of food for juvenile fish and crabs (Chao and Musick, 1977; Bell and Coull, 1978; Holland *et al.*, 1989). Estuarine benthos also have important roles in ecological processes that affect water quality and productivity. For example, the feeding and burrowing activities of macrobenthos affect sediment depositional patterns and chemical transformations (Carriker, 1967; Rhoads, 1974; Kemp and Boynton, 1981). Benthic feeding activities can remove large amounts of particulate materials from shallow estuaries, which may improve water clarity (Cloern, 1982; Officer *et al.*, 1982; Holland *et al.*, 1989).

A benthic index was developed using the Virginian Province 1990 Demonstration Project data. The procedures used to develop the Benthic Index in 1990 are documented in the EMAP-Estuaries Virginian Province 1990 Demonstration Project Report (Weisberg *et. al.*, 1993). This index is based on 5 parameters extracted from the raw Benthic Infaunal data base for each EMAP base station.

The "discriminate score" was calculated using transformed and normalized scores according to the equation below:

Discriminate score =

*0.01053 * percent Expected species {mean number}*

\+ *0.81692 * Number of Amphipods*

\+ *0.57697 * Average weight for each polychaete*

\+ *0.46526 * Number of Capitellid worms*

\+ *0.67136 * Percent of total abundance which are bivalves*

A "critical" value of 3.4 was determined as the delimiter between degraded conditions (BI $\leq$ 3.4) and non-degraded conditions (BI > 3.4).

The "Expected number of species" at each station was calculated based on a habitat gradient established in 1990. This gradient indicated that the number of species found at a station is related to the salinity of the bottom waters at the station. Given similar anthropogenic influences at two stations, the number of species found would be strongly influenced by the salinity of the bottom waters. In order to normalize the expected number of species to the salinity habitat gradient, the baseline (or expected number of species) was established by fitting a polynomial regression of the 90th percentile of a 3 part per thousand (salinity) moving average of species richness versus salinity (Weisberg *et al.*, 1993).

Table B-1. Results of benthic index validation exercise. Numbers represent the number of stations in that cell of the matrix. BI = benthic index score.

	Reference Stations	Degraded Stations	
BI>3.4 *(Reference)*	39 (85%)	6 (46%)	45
BI< 3.4 *(Degraded)*	7 (15%)	7 (54%)	14
	46	13	59

B.2 VALIDATION OF THE 1990 BENTHIC INDEX

The process of validation involved testing the 1990 benthic index with an independent data set to ensure that the multivariate solution was not specific to the original test data set used in 1990. Using the same criteria applied in 1990 to define "degraded" and "undegraded" (*i.e.*, reference) stations, an independent data set was established from the 1991 database. This test data set consisted of 13 stations classified as degraded based on high sediment contaminant concentrations (combined with toxicity) and/or low near-bottom dissolved oxygen levels, and 46 stations classified as reference based on low sediment contaminant levels and the absence of toxicity and low dissolved oxygen conditions.

Of the 46 stations from 1991 classified as "reference", the 1990 benthic index correctly classified 39 (85%) and misclassified 7 (15%) (Table B-1). Of the 13 stations from 1991 classified as "degraded", the 1990 benthic index correctly classified 7 (54%) and misclassified 6 (46%). For these degraded stations, the rate of correct classification was much higher for low dissolved oxygen stations versus stations with contaminated sediments. The relatively high overall rate of misclassification, particularly for degraded stations, was deemed unacceptable, and a decision was made to reconstruct a new benthic index using the combined 1990 and 1991 data sets.

B.3 RECONSTRUCTION OF THE BENTHIC INDEX USING 1990-91 DATA

Reconstruction of the benthic index using the combined 1990 and 1991 data followed the same basic steps described in the 1990 Demonstration Project Report (Weisberg *et. al.* 1993). Results and discussion are presented in the following sections.

Step 1: Develop a test data set

The following criteria (from 1990) were used to select stations for reconstructing the benthic index:

DEGRADED STATIONS:

1.) Stations exhibiting sediment toxicity (*i.e.*, % survival less than 75 and significantly different from controls) and having one or more contaminants exceeding Long and Morgan (1990) ER-M values.

OR

2.) Stations exhibiting DO below 0.3 mg/L at any time, or 10% of continuous measurements less than 1 mg/L, or 20% less than 2 mg/L, or less than 2 mg/L for 24 consecutive hours.

REFERENCE STATIONS:

1.) Stations where no contaminant exceeded the ER-M value, no sediment toxicity was observed (*i.e.*, % survival greater than 75% and not significantly different from controls), and for which bottom DO was never less than 1 mg/L, 90% of the continuous DO measurements were greater 3 mg/L and 75% of the DO measurements were greater than 4 mg/L. (NOTE: because of this criteria, only those 1991 stations having continuous DO records were considered as potential reference stations).

Once an initial set of stations was identified using this set of criteria, the list was reviewed carefully to eliminate any reference sites located in areas potentially subject to physical disturbance, such as dredged shipping channels. Four sites were eliminated through this process. The final test data set, combining 1990 and 1991 data, consisted of 31 degraded stations and 53 reference stations (Tables B-2 and B-3). Both the degraded and reference stations encompassed a wide range of habitats, including all major salinity zones and sediment types that occur in the Virginian Province.

Step 2: Identify candidate benthic measures

As in 1990, benthic abundance, biomass, and species composition data were used to define 28 descriptors of the major ecological attributes of the benthic assemblages occurring at each sample site (Table B-4). In constrast to 1990, data for epifauna (organisms that live on hard surfaces such as shells) were included in the 1991 computations in an attempt to determine whether any of the resultant metrics could discriminate reliably between reference and degraded conditions.

Estuaries are characterized by large natural variations in certain physiochemical conditions (*e.g.*, salinity, sediment grain size) known to be major factors controlling the diversity and abundance of resident biota. Such factors must be identified and controlled for before the responses of candidate benthic measures to pollution exposure can be characterized accurately. Pearson correlation coefficients were calculated to determine relationships between the individual metrics listed in Table B-4 and various physical habitat variables including sediment silt-clay content, salinity, water depth, latitude, and sediment total organic carbon (TOC) concentration. These relationships were determined using data from the reference stations only to avoid the potentially confounding effects of stressors (*e.g.*, contaminants and low dissolved oxygen) known to occur at the degraded sites.

Many of the candidate benthic measures were significantly ($p<0.05$) correlated with at least one of the habitat factors measured (Table B-5). However, only five of the correlations accounted for a significant proportion of the total variation, defined here as more than 25%. Four of these five were measures of species richness (both total number of infaunal species per event and mean number of infaunal species per grab), which were both positively correlated with bottom salinity and negatively correlated with TOC. The fifth correlation was a positive one between the mean abundance of equilibrium species and bottom salinity. Relationships between the rest of the candidate measures and the other habitat factors (*i.e.*, latitude, silt-clay content of sediments, and water depth) occurred less frequently and did not account for as much of the total variation as relationships with salinity and TOC (Table B-5).

A three dimensional plot of the mean number of benthic infaunal species per grab versus salinity and TOC was generated, and a quadratic response surface model was fit to predict the "expected" number of infaunal species:

Expected number of species =

$$
\begin{aligned}
& 8.25 + (0.000387\ (TOC)) \\
& - (1.9\times10^{-8})(TOC))^2) \\
& + (0.784\ (salinity)) - (0.00125\ (salinity)^2) \\
& - (0.00002031\ (TOC)\ (salinity)).
\end{aligned}
$$

The "expected" number of species calculated from this relationship ($r^2 = 0.61$) is presumed to represent the baseline (*i.e.*, reference) response of benthos to estuarine gradients in salinity and TOC in the absence of known stressors (*e.g.*, chemical contaminants and low dissolved oxygen conditions). An adjusted measure of species richness was then determined for each station by calculating the percent deviation in the mean number of infaunal species per grab from this baseline condition. This adjusted measurement was termed the percent expected number of species:

$$= \frac{number\ of\ species\ present}{expected\ number\ of\ species} \times 100$$

Table B-2. List of degraded stations from 1990 or 1991 used in reconstructing the benthic index.

Habitat Class	Degraded Stations		
	Low Dissolved Oxygen Stress	*Contaminant Stress*	*Low Dissolved Oxygen and Contaminant Stress*
Low salinity (< 5 ppt)	*None*	*Houstatonic River (169)* 41°17'12" 73°04'19" *Delaware River (223)* 39°45'00" 75°29'00" *Susquehanna River (351)* 39°34'41" 76°05'29" *Delaware River (365)* 40°06'04" 74°50'11"	*Anacostia River (088)* 38°52'11" 76°59'51"
Brackish (5-18 ppt)	*Chesapeake Bay (062)* 38°59'12" 76°21'29" *Chesapeake Bay (065)* 38°33'27" 76°24'05" *Potomac River (180)* 38°04'13" 76°27'53" *Potomac River (182)* 38°13'06" 76°47'08" *Breton Bay (306)* 38°13'41" 76°41'48" *Breton Bay (312)* 38°15'22" 76°39'39" *Chesapeake Bay (325)* 38°37'29" 76°27'52"	*Colgate Cove (082)* 39°15'12" 76°33'06"	*Bear Creek (081)* 39°14'36" 76°29'29" *Patapsco River (134)* 39°14'47" 76°33'25" *Passaic River (103)* 40°45'00" 74°09'54"
Estuarine (>18 ppt)	*Chesapeake Bay (282)* 37°39'02" 76°12'52" *Chesapeake Bay (056)* 38°08'40" 76°14'05" *Chesapeake Bay (080)* 38°53'23" 76°24'04"	*Arthur Kill (094)* 40°37'18" 74°12'12" *Raritan Bay (260)* 40°42'17" 74°06'59" *Raritan River (369)* 40°30'40" 74°18'00" *Kill Van Kull (373)* 40°38'52" 74°04'29" *Flushing Bay (375)* 40°46'37" 73°51'19" *Flushing Bay (377)* 40°47'31" 73°51'41" *East River (378)* 40°47'31" 73°55'54" *Taunton River (419)* 41°42'38" 71°09'49" *Taunton River (421)* 41°46'00" 71°07'23"	*Elizabeth River (086)* 36°49'55" 76°17'38" *Blackrock Harbor (098)* 41°09'35" 73°12'37" *New Bedford Harbor (099)* 41°38'33" 70°54'42"

Table B-3. List of reference stations from 1990 and 1991 used to reconstruct the benthic index. Data from both 1990 and 1991 were used for LTT stations 188 and 150.

Habitat Class	Reference Stations		
Low Salinity (< 5 ppt)	Hudson River (101) 41°23'00" 73°57'23"	Potomac River (188) 38°44'12" 77°02'00"	Hudson River (424) 42°08'00" 73°54'26"
	Rappahannock River (192) 37°57'54" 76°52'30"	Delaware River (358) 39°50'05" 75°21'03"	Potomac River (333) 38°51'32" 77°02'03"
	Elk River (254) 39°28'47" 75°56'30"	James River (273) 37°14'26" 76°57'18"	Potomac River (326) 38°37'30" 77°09'43"
Brackish (5-18 ppt)	James River (269) 37°09'48" 76°37'43"	Wye River (336) 38°54'22" 76°10'15"	Honga River (316) 38°18'13" 76°11'01"
	Chesapeake Bay (058) 39°07'45" 76°16'53"	Fishing Bay (317) 38°18'55" 76°01'11'	Chesapeake Bay (057) 37°26'04" 76°14'07"
	Chesapeake Bay (339) 39°03'14" 76°25'16"	Little Choptank River (322) 38°31'10" 76°16'07"	Chesapeake Bay (307) 38°13'41" 76°05'18"
	St. Clements Bay (314) 38°16'53" 76°42'39"	Tangier Sound (041) 38°01'41" 75°54'06"	Delaware Bay (344) 39°16'41" 75°16'28"
	Rappahannock River (288) 37°49'53" 76°44'49"	Chesapeake Bay (291) 37°55'53" 76°15'22"	
Estuarine (> 18 ppt)	Barnegat Bay (256) 39°56'36" 74°06'07"	Chesapeake Bay (265) 37°05'20" 76°07'54"	Narragansett Bay (070) 41°38'29" 71°18'01"
	Tangier Sound (285) 37°48'53" 75°55'25"	Long Island Sound (386) 41°05'26" 73°05'04"	Nantucket Sound (402) 41°18'56" 70°07'05"
	Delaware Bay (342) 39°09'30" 74°56'16"	Jamaica Bay (371) 40°36'57" 73°53'15"	Westport River (413) 41°31'49" 71°05'40"
	Back River (266) 37°05'53" 76°20'00"	Long Island Sound (388) 41°06'30" 72°38'43"	Vineyard Sound (407) 41°26'46" 70°40'43"
	Chesapeake Bay (283) 37°39'59" 76°00'26"	Rehobeth Bay (328) 38°40'29" 75°07'05"	Buzzards Bay (406) 41°26'30" 70°54'00"
	Chesapeake Bay (270) 37°13'16" 76°15'19"	Indian River Bay (150) 38°35'36" 75°06'42"	Nantucket Sound (408) 41°27'01" 70°27'25"
	Chincoteague Bay (305) 38°13'24" 75°13'02"	Niantic River (405) 41°20'35" 72°10'45"	Buzzards Bay (414) 41°35'02" 70°47'48"
	Great Bay (349) 39°30'13" 74°23'05"	Napeague Bay (159) 41°03'42" 72°00'06"	Nantucket Harbor (404) 41°20'00" 70°01'00"
	Delaware Bay (338) 39°00'40" 75°01'34"	Delaware Bay (335) 38°51'50" 75°06'50"	
	Delaware Bay (033) 39°12'36" 75°12'43"	Block Island Sound (390) 41°07'53" 71°59'08"	

Table B-4. Candidate benthic measures used to formulate the benthic index. T-tests were used to test equality of the means for each metric for degraded versus reference (nondegraded) sites.

Candidate metrics	t-test (p-value)	Direction (+ = greater mean value at reference sites)
Measures of Biodiversity		
Shannon-Weiner index	< 0.001	+
Proportion of expected # of species	< 0.001	+
Measures of Community Condition		
Total benthic biomass per station	0.87	+
Mean biomass per grab	0.87	+
Total infaunal abundance for station	0.48	-
Mean infaunal abundance per grab	0.48	-
Total # of infaunal species/station	< 0.001	+
Mean # of infaunal species/grab	< 0.001	+
Total # of epifaunal species/station	< 0.001	+
Mean epifaunal abundance per grab	0.07	+
Mean # of epifaunal species/grab	< 0.001	+
Measures of Individual Health		
Biomass/abundance ratio	0.04	+
Weight per individual polychaete	0.84	-
Weight per individual mollusc	0.10	+
Weight per individual bivalve	0.06	+
Measures of Functional Groups		
Suspension feeding species abundance	0.55	-
Deposit feeding species abundance	0.37	-
Omnivore/predator species abundance	< 0.001	+
Opportunistic species abundance	0.06	-
Equilibrium species abundance	0.001	+
Measures of Taxonomic Composition		
Amphipod abundance	0.87	-
Bivalve abundance	0.32	-
Gastropod abundance	< 0.001	+
Molluscan abundance	0.43	-
Polychaete abundance	0.22	+
Capitellid polychaete abundance	0.34	+
Spionid polychaete abundance	0.28	-
Tubificid oligochaete abundance	0.34	-

Table B-5. Summary of correlation between habitat indicators and the candidate benthic measures.

Habitat Indicator	Number of Significant Correlations	Number of Correlations with $r^2 > 0.10$	Number of Correlations with $r^2 > 0.25$
Salinity (ppt)	16	15	3
Latitude	4	2	0
Silt-clay Content	8	7	0
Water Depth	8	4	0
Total organic carbon	10	8	2

Step 3: Identify combinations of candidate benthic measures that discriminate between degraded and nondegraded areas

A series of discriminant analyses were run in succession to identify the benthic measures from Table B-4 which best discriminated between the degraded and nondegraded stations in the test data set. In some instances, the variable selection process included an initial t-test (assuming unequal variances) to eliminate any variables not having significantly different means (t-test at $p < 0.2$) at degraded versus nondegraded stations (Table B-4). In other instances, the t-test was not performed, and all variables were used in a stepwise discriminant analysis. In addition, discriminant analyses were performed using both untransformed and selected transformed variables (*e.g.*, $\log_{10}$ transformed species abundance variables), as well as both unadjusted and adjusted measures of species richness (adjusted for the effects of salinity and TOC).

The results of the various discriminant analyses are summarized in Table B-6. Three variables were included in the model generated by the first stepwise discriminant analysis (Index 1): 1) mean abundance of opportunistic species, 2) biomass/abundance ratio for all species collected at a site, and 3) the mean number of infaunal species. This combination of measures correctly classified 84% of the degraded sites and 85% of the nondegraded reference sites (Table B-6), using the cross-validation procedure of the discriminant analysis. The canonical r^2, which approximates the total variance explained by this analysis, was 0.40.

Certain variables were eliminated as a result of performing the t-test prior to running the first stepwise discriminant analysis. When the same analysis was run without first performing a t-test, two additional measures were selected (Index 2): 1) weight of individual polychaetes, and 2) mean abundance of capitellids. This combination of measures resulted in a slightly better canonical r^2 (0.44), but the rate of correct classification for degraded stations decreased to 77% (two fewer stations correctly classified) while the rate for reference stations remained unchanged.

In the third discriminant analysis, salinity and TOC-adjusted measures of species richness (percent of the expected mean number of species) were used in place of the unadjusted values. Five measures were selected in this model (Index 3): 1) Shannon-Weiner diversity index, 2) mean abundance of spionids, 3) mean abundance of equilibrium species, 4) mean abundance of tubificids, and 5) mean weight of individual bivalves. This model correctly classified 61% of the degraded stations and 89% of the reference stations, with a canonical r^2 of 0.43.

The failure of the adjusted species richness measure (TOC/salinity adjustement) to appear in indices 1 and 2, indicates that it was less effective at discriminating between degraded and reference conditions than the unadjusted measure. To verify this, the percent of expected mean number of species was forced into Index 1 in place of the unadjusted mean number of infaunal species. This resulted in a lower rate of correct classification for degraded stations (74%) and a marginally improved rate of classification for reference stations (87%) compared to the original Index 1. The canonical r^2 for this model was 0.42. The

Table B-6. Results of discriminant analyses conducted to combine candidate benthic measures into an index.

		Cross-validation Efficiency		
Analysis	**Selected Measures**	**Percent of Degraded Sites Correctly Classified**	**Percent of Reference Sites Correctly Classified**	**Canonical R^2**
Index 1 • t-test on raw variables • step-wise discriminant analysis • no variables forced	• Mean abundance of opportunistic species • Biomass/abundance ratio, all species • Mean number of infaunal species	84%	85%	0.40
Index 2 • no t-test • raw variables • step-wise discriminant analysis • no variables forced	• Mean abundance of opportunistic species • Biomass/abundance ratio, all species • Mean number of infaunal species • Weight of individual polychaetes • Mean abundance of capitellids	77%	85%	0.44
Index 3 • species richness adjusted for salinity/TOC • unadjusted species richness variables removed • no t-test • step-wise discriminant analysis • no variables forced • raw variables used	• Shannon-Weiner diversity index • Mean abundance of spionids • Mean abundance of equilibrium species • Mean abundance of tubificids • Mean weight of individual bivalves	61%	89%	0.43
Index 4 • no t-test • $\log_{10}$ transform all species abundance variables except total infaunal abundance • step-wise discriminant analysis • no variables forced	• Mean amphipod abundance • Mean abundance of equilibrium species • Mean weight of individual polychaetes • Total infaunal abundance	77%	91%	0.59
Index 5 • $\log_{10}$ transform all variables • no t-test • step-wise discriminant analysis • no variables forced	• Mean abundance of opportunistic species • Mean abundance of amphipods • Mean abundance of tubificids • Mean weight of individual bivalves • Mean weight of individual molluscs • Mean number of infaunal species	67%	92%	0.53

net result of using the adjusted species richness measure in place of the unadjusted one was to misclassify an additional degraded site and correctly classify one additional reference site, with an almost negligible increase in the canonical r^2.

During the development of the benthic index in 1990, a $\log_{10}$ (value + 1) transformation was applied to some of the measures to make their distributions normal and variances homogenous. Although transforming individual variables to achieve normality does not imply that the multivariate distribution used in discriminant analysis will be "multivariate normal", the $\log_{10}$ (value + 1) transformation was applied, with one exception, to all species abundance data prior to the analysis for Index 4 (Table B-6). Due to an oversight, the log transformation was not applied to the total infaunal abundance measure. Four measures were selected in this model (Index 4): 1) mean amphipod abundance, 2) mean abundance of equilibrium species, 3) mean weight of individual polychaetes, and 4) total infaunal abundance. This model correctly classified 77% of the degraded stations and 91% of the reference stations, with a canonical r^2 of 0.59.

The model used in Index 4 explained significantly more of the total variation than any other model (*i.e.*, highest canonical r^2), and it correctly classified three more reference stations, but two fewer degraded stations, compared to Index 1. When the mistake discussed above was corrected and all abundance variables were log-transformed, the resultant model included the following variables: 1) Shannon-Weiner diversity index, 2) mean abundance of amphipods, 3) mean abundance of opportunistic species, 4) total infaunal abundance, 5) mean weight of individual bivalves, and 6) mean abundance of gastropods. This "corrected" model correctly classified 74% of degraded stations and 87% of reference stations, with a canonical r^2 of 0.50.

For Index 5 (Table B-6), all of the candidate variables were $\log_{10}$ (value + 1) transformed and the following were selected: 1) mean abundance of opportunistic species, 2) mean abundance of amphipods, 3) mean abundance of tubificids, 4) mean weight of individual bivalves, 5) mean weight of individual molluscs, and 6) mean number of infaunal species. This model correctly classified 67% of degraded stations and 92% of the reference stations, with a canonical r^2 of 0.53.

Step 4: Validating the model

Five potential indices were examined in step 3. The Virginian Province benthic data from the 1992 sampling season, which represents an independent data set, will be used in the future to provide a true validation of each of these indices. In the absence of this independent validation, the discriminant analysis cross-validation procedure was used to determine the rates of correct classification presented in Table B-6. In this process one individual station is removed from the test dataset, the indices recalculated, and the station removed from the test dataset reclassified according to the new index. This process is repeated for each station in the original test dataset. The overall rate of correct classification, based on the cross-validation procedure, were 85%, 82%, 79%, 86% and 83% for the five candidate benthic indices, respectively.

Step 5: Scaling the index

The final step in developing the benthic index was to select one of the indices based on the calibration and validation information, calculate canonical discriminant scores for all sample sites, and adjust the scores to result in a critical value of zero. Although all five candidate indices were cross-validated at an acceptable level (i.e., about 80% overall correct classification), the first alternative (Index 1 in Table B-6) was used to assess the status of benthic resources of the Virginian Province for the combined 1990-91 data set. The discriminant function for this index was:

Discriminant Score =

*-0.68 * Mean abundance of opportunistic species*

*+ 0.36 * Biomass/abundance ratio for all species*

*+ 1.14 * Mean number infaunal species per grab.*

When applied to all 392 sites sampled in the Virginian Province in 1990 and 1991, the range in canonical discriminant scores for this index was -7.78 to +6.33, with a critical value for discriminating between degraded and reference sites of -0.5 (calculated as the point giving the optimal correct classification efficiency for both reference and degraded sites). A value of 0.5 was then added to all scores to result in a critical value of zero, *i.e.*, a negative score indicated degraded conditions. An offset was selected in place of a scaling factor (*i.e.*, scaling from 0 to 10), because a scaling factor requires recalculation every year resulting in a new critical value each year. An offset is not affected by the range of values, therefore, the critical value will remain constant between years.

B.4 DISCUSSION OF THE BENTHIC INDEX

Although only five candidate indices were presented in Table B-6, it is important to note that many variations of these indices were examined in an attempt to find the best model. These efforts revealed that for the 1990-91 data set, optimization of the model represented a "trade-off" among three competing factors: 1) classification efficiency, 2) statistical robustness (as expressed by the canonical r^2), and 3) complexity/ecological relevance. Essentially, it was found that increases in the canonical r^2, which reflect increased ability of the model to explain variation, were accompanied by concomitant decreases in classification efficiency and increases in the complexity of the model. Complexity is defined here to be a direct function of the number of variables used, with models based on fewer variables being more desirable because presumably they are easier to comprehend and appreciate. The concept of "ecological relevance" also is important in this context, in that the variables chosen for the index should conform at least to some degree with currently accepted scientific knowledge about the response of benthic communities to stress. A quick check on whether a discriminant function is plausible is to examine the signs of the coefficients. The sign on a coefficient should make sense in ecological terms, *i.e.*, we would expect to find a greater number of species at a reference station than at a degraded site, therefore, we would expect a "+" coefficient.

Although Index 4 had the best overall rate of correct classification, it correctly classified fewer degraded stations than Index 1 and is more difficult to justify in ecological terms. Although it had the lowest canonical r^2 value, Index 1 had the second best overall classification efficiency and the highest rate of correct classification for degraded stations. This index also is the least complex of all the candidate indices, and was appealing from an "ecological relevance" standpoint in that it is comprised of one measure of individual health (biomass/abundance ratio), one measure of functionality (abundance of opportunists), and one measure of community condition (mean number of infaunal species). The signs of the discriminant coefficients are also ecologically interpretable for Index 1. Higher values for mean number of infaunal species and the biomass/abundance ratio can be defended as being associated with healthy conditions (positive coefficient), and an increase in the abundance of opportunistic organisms at degraded sites (negative coefficient) is consistent with ecological theory.

Compared to Index 1, the second, third and fifth indices had better canonical r^2 values and better rates of classification for reference stations. This suggests that they were better able to explain the variation present at the reference sites. However, all three indices had slightly worse classification efficiencies for the degraded stations and were more complex. Therefore, Index 1 was selected as the best index.

APPENDIX C

SUB-POPULATION ESTIMATES FOR CHESAPEAKE BAY AND LONG ISLAND SOUND

The two largest systems within the Virginian Province are Chesapeake Bay (11,469 km^2) and Long Island Sound (3,344 km^2). Combined these two systems represent 63% of the surface area of the entire Province. Because of their size, and therefore the number of sampling locations in each, estimates of ecological condition of these systems are possible using the EMAP design. However, the level of uncertainty will remain higher than for estimates for the Province as a whole or individual classes.

This appendix provides the tools for generating these estimates, *i.e.*, data for these two systems are summarized using CDFs and bar charts. Each system is defined as including all adjacent tributaries and small systems. For example, the data set for Chesapeake Bay includes the Potomac, James, and Rappahannock Rivers, and all the small systems connecting to the mainstem of the Bay. Since the Long Island Sound data set contains no large tidal rivers and fewer small systems than Chesapeake Bay, this may account for some of the differences observed between these two systems. Fifty one stations are included in the Chesapeake Bay data set and 14 in the Long Island Sound data set.

C.1 Biotic Condition Indicators

C.1.1 Benthic Index

CDFs for the benthic index are illustrated in Figure C-1. Approximately 16 ± 10% of the sampled area of Chesapeake Bay produced a benthic index value below zero, and the corresponding area of Long Island Sound was 14 ± 21%.

C.1.2 Number of Benthic Species

The mean number of species collected per grab at each station, as percent area in these systems, is illustrated in Figure C-2. The distribution and maximum (42 and 44 species) values are similar for Chesapeake Bay and Long Island Sound, respectively.

C.1.3 Total Benthic Infauna Abundance

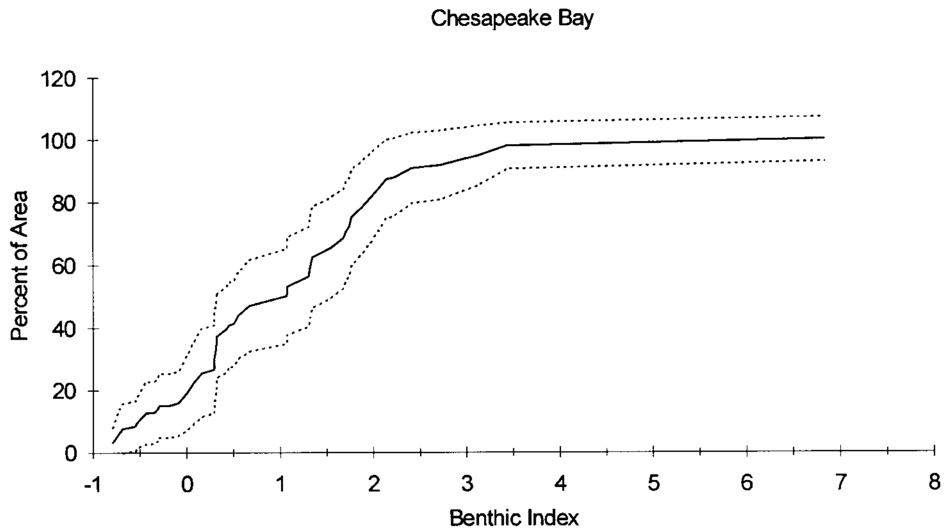
Figure C-3 shows the distribution of total number of benthic individuals per m^2 measured in Chesapeake Bay and Long Island Sound. The maximum number of individuals collected at a station was higher in the Sound than in the Bay (13,992 and 8,561, respectively).

C.1.4 Number of Fish Species

The number of fish species collected per standard trawl is shown in Figure C-4. Between 0 and 7 species the distributions are similar, however the maximum number if individuals caught at a station was approximately double in Chesapeake Bay what it was in Long Island Sound (15 and 7, respectively).

C.1.5 Total Finfish Abundance

The total number of fish captured per standard trawl (catch per unit effort) was greater at Chesapeake Bay stations than Long Island Sound stations (Figure C-5). The maximum catch in the Bay was 650 individuals, whereas, no more than 69 were collected at any station in Long Island Sound. This is presumably due to habitat and cannot be related to man's impact.



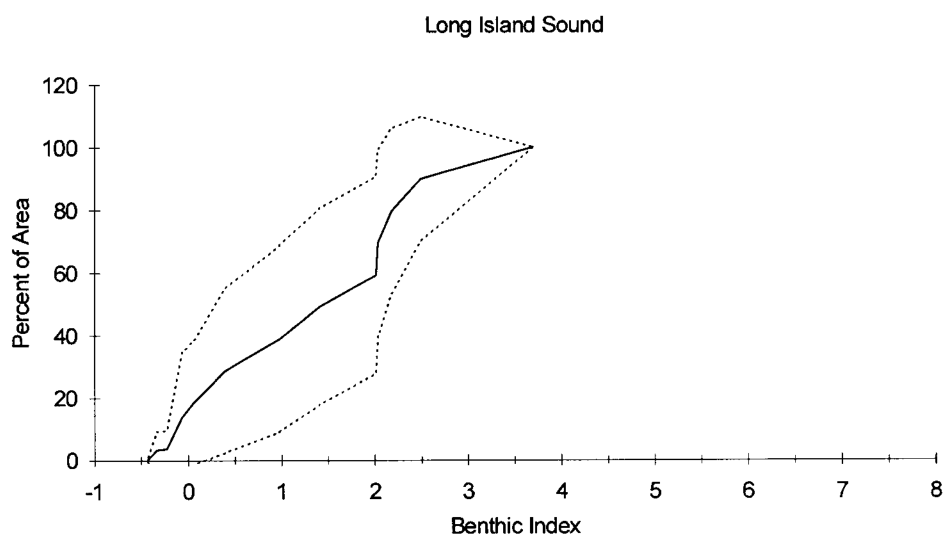
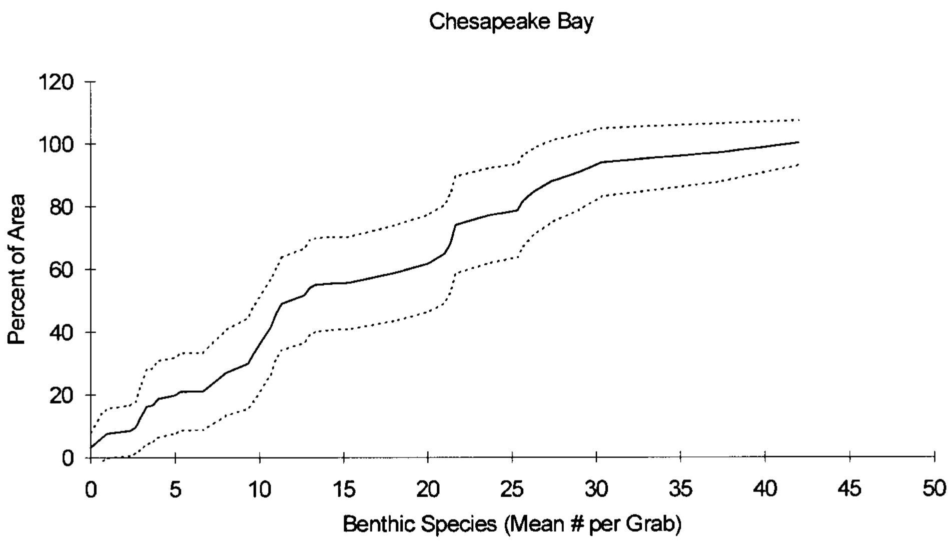


Figure C-1. Cumulative distribution functions of benthic index as a percent of area of Chesapeake Bay and Long Island Sound, 1991. (Dashed lines are the 95% confidence intervals).



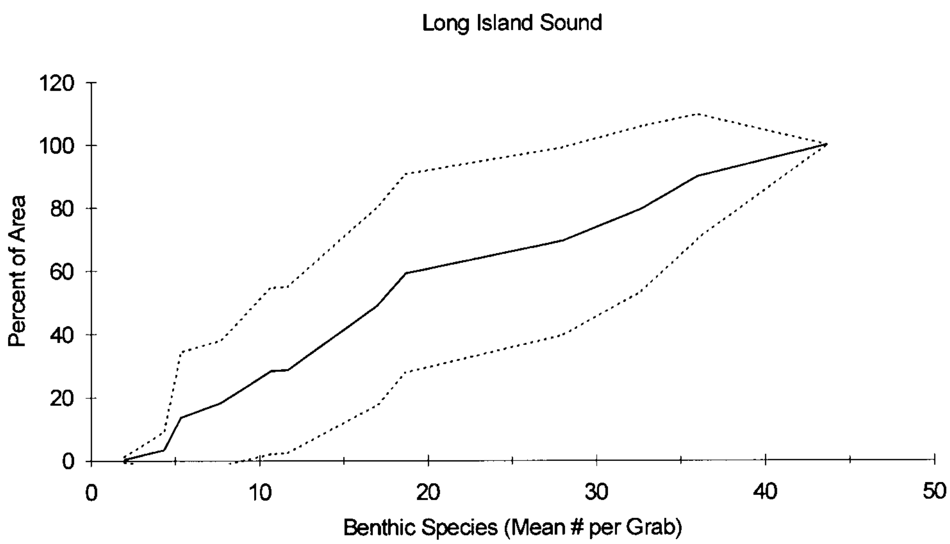
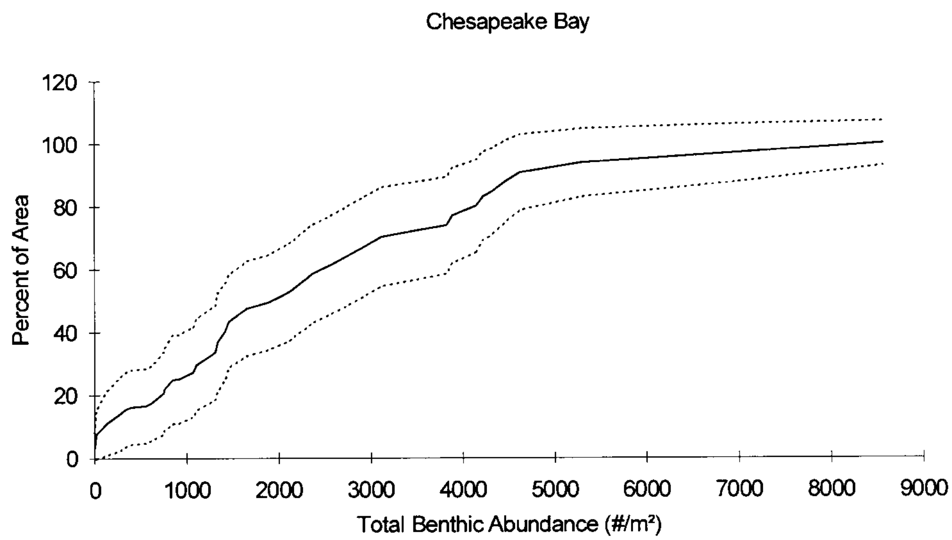


Figure C-2. Cumulative distribution functions of the mean number of benthic invertebrate species collected per grab as a percent of area of Chesapeake Bay and Long Island Sound, 1991. (Dashed lines are the 95% confidence intervals).



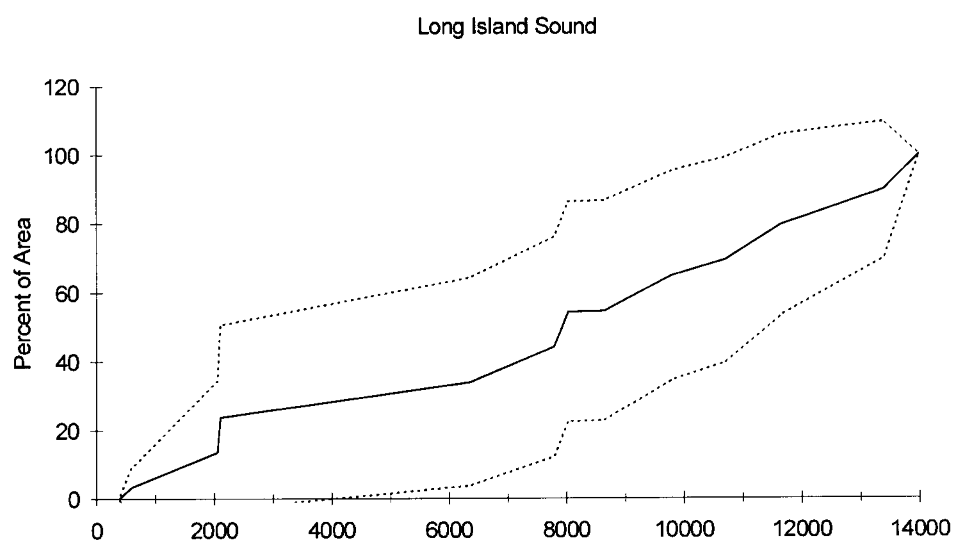


Figure C-3. Cumulative distribution functions of the number of benthic invertebrates collected per m^2 as a percent of area of Chesapeake Bay and Long Island Sound, 1991. (Dashed lines are the 95% confidence intervals).

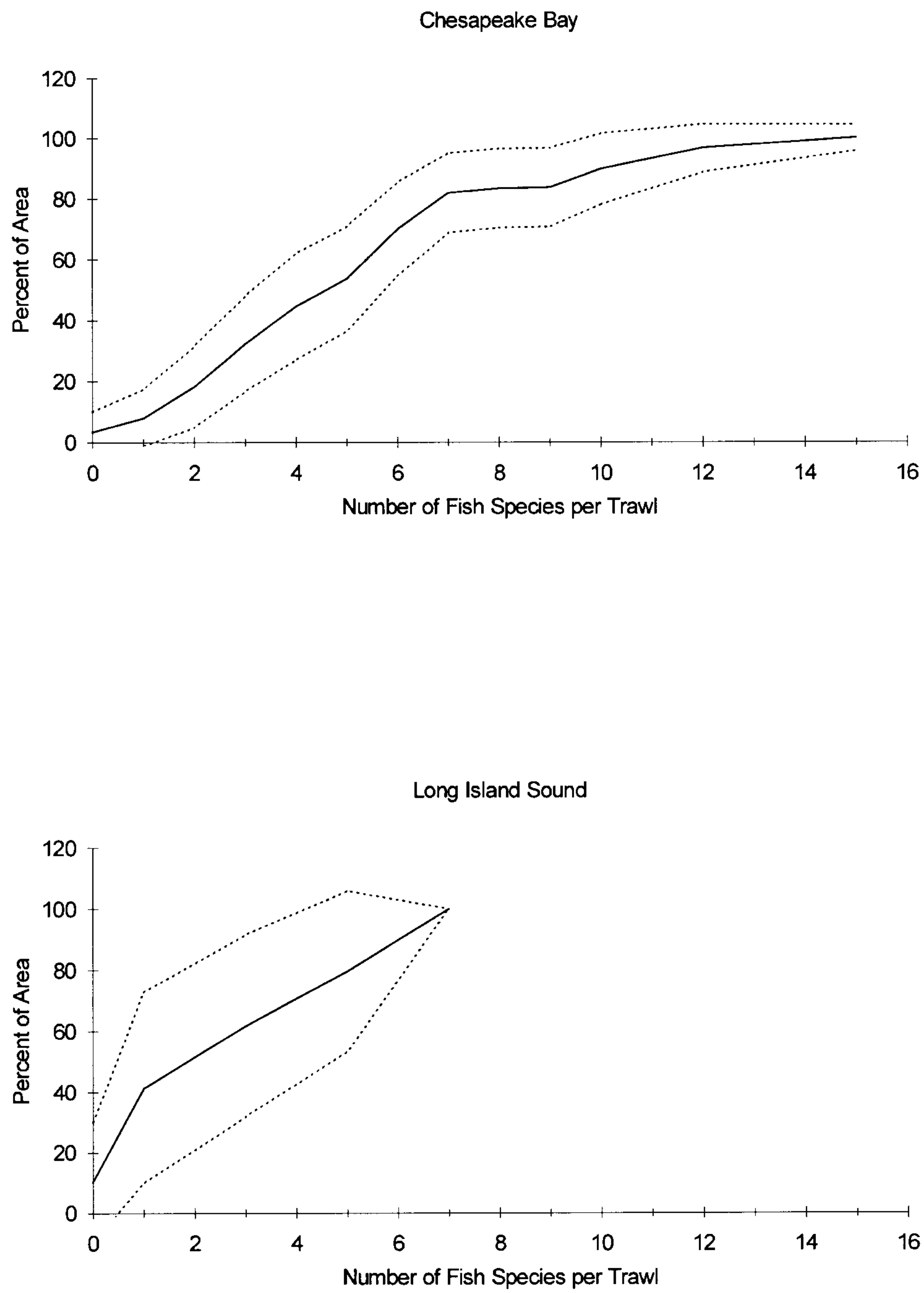
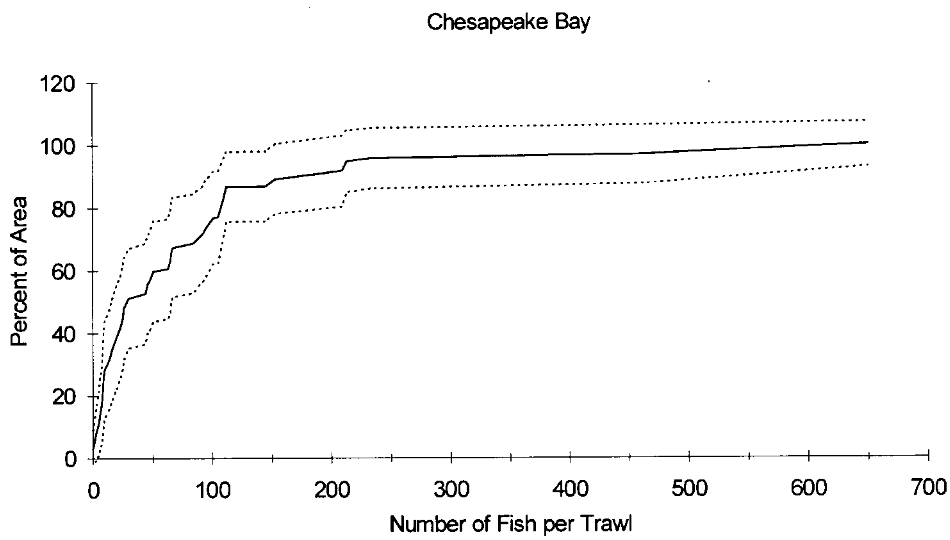


Figure C-4. Cumulative distribution functions of the number of fish species collected in standard trawls as a percent of area of Chesapeake Bay and Long Island Sound, 1991. (Dashed lines are the 95% confidence intervals).



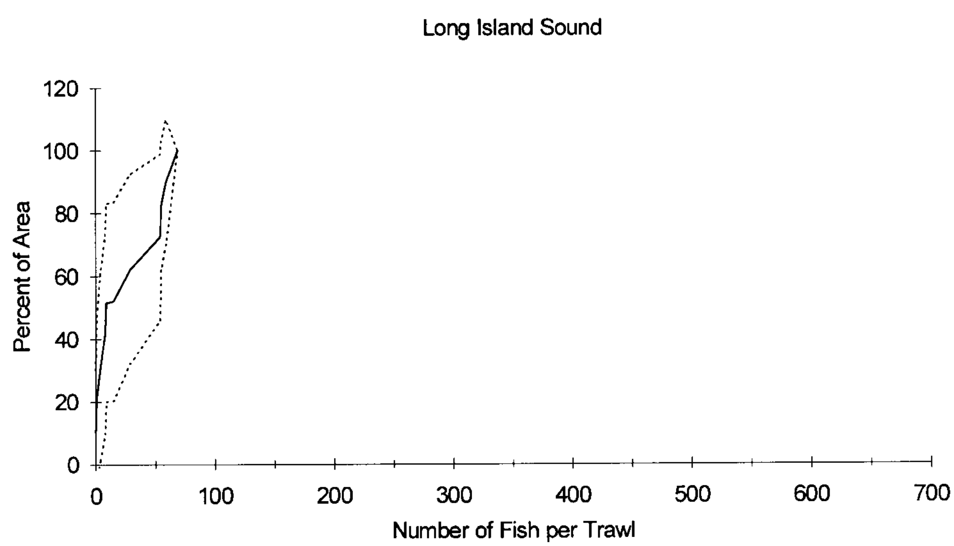


Figure C-5. Cumulative distribution functions of the number of fish caught per standard trawl as a percent of area of Chesapeake Bay and Long Island Sound, 1991. (Dashed lines are the 95% confidence intervals).

C.1.6 Fish Gross External Pathology

All target species were examined for evidence of gross external pathologies. The rate of pathologies in Chesapeake Bay and Long Island Sound were similar, however, only 159 fish were collected and examined in Long Island Sound compared to 1,265 in Chesapeake Bay (Table C-1).

C.1.7 Fish Tissue Contamination

As discussed in Section 3, no composite samples analyzed contained organic contaminants in excess of stated action limits. Of the 41 composites analyzed from Chesapeake Bay, 10% exceeded the mean international action limit of 2 μg/g wet weight for arsenic. All four composites were spot and atlantic croaker. In Long Island Sound, 44% of the nine composites exceeded the arsenic limit, and all six composites were winter flounder.

C.2 Abiotic Condition Indicators

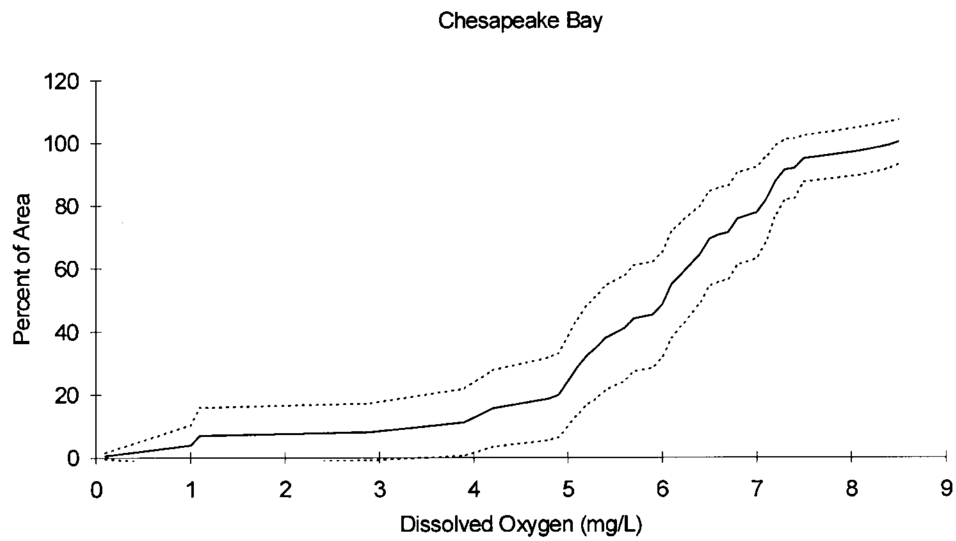
C.2.1 Instantaneous Dissolved Oxygen Concentration

CDFs for bottom dissolved oxygen concentration in Chesapeake Bay and Long Island Sound are shown in Figure C-6. The percent of sampled area of Chesapeake Bay with severely hypoxic water (DO ≤2 mg/l) was similar that of Long Island Sound (7 ± 9% and 10 ± 20%, respectively). Approximately 44 ± 32% of the Sound was marginal, with DO values less than 5 mg/L (compared to 20 ± 13% for the Bay).

Table C-1. Incidence of gross external pathology for Chesapeake Bay and Long Island Sound observed by field crews among target fish species in 1991.

	Lumps	Growths	Ulcers	Fin Rot	Total
Chesapeake Bay					
Frequency	1	0	9	2	11[a]
Total # Fish Examined	1,265	1,265	1,265	1,265	1,265
Percent Incidence	0.08%	0%	0.71%	0.16%	0.87%
Number Stations Represented					4
Long Island Sound					
Frequency	0	0	0	1	1
Total # Fish Examined	159	159	159	159	159
Percent Incidence	0%	0%	0%	0.63%	0.63%
Number Stations Represented					1

[a] one fish was found with two pathologies, therefore 12 pathologies were identified on 11 fish.



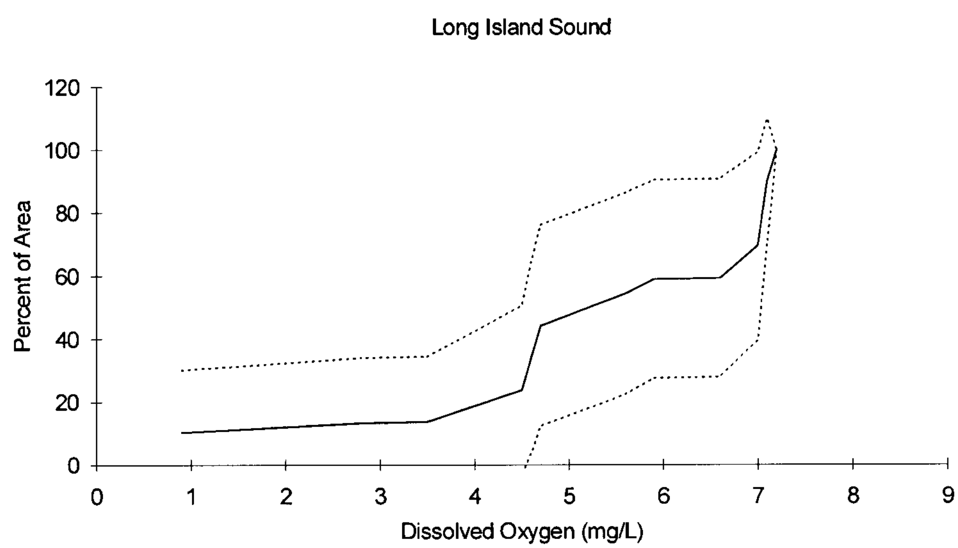


Figure C-6. Cumulative distribution functions of bottom dissolved oxygen concentration as a percent of area of Chesapeake Bay and Long Island Sound, 1991. (Dashed lines are the 95% confidence intervals).

C.2.2 Dissolved Oxygen Stratification

The difference in measured DO concentrations at the bottom compared with surface measurements taken at those same stations are illustrated in Figure C-7. No significant differences are shown by these graphs.

C.2.3 Sediment Toxicity

Sediments were classified as toxic if amphipod survival in the test sediment was less than 80% of that in the control sediment, and significantly different. Approximately 20 ± 13% and 23 ± 30% of the area of Chesapeake Bay and Long Island Sound respectively exhibited toxic sediments (Figure C-8).

C.2.4 Sediment Contaminants - Organics

Draft EPA Sediment Quality Criteria (SQC) exist for four compounds for which EMAP is monitoring: Acenaphthene, phenanthrene, fluoranthene, and dieldrin. One station at the mouth of Chesapeake Bay exceeded the SQC for all three PAHs; however, as discussed in Section 3, this is believed to be due to contamination of the sample. No other station in Chesapeake Bay or Long Island Sound exceeded any of the SQCs.

CDFs for combined PAHs are presented in Figure C-9. Although the maximum concentration measured was higher in Long Island Sound than Chesapeake Bay (22,700 and 4,091 ng/g dry weight, respectively), the distributions are similar with 96 ± 6% of the sampled area of Long Island Sound containing concentrations less than 4,000 ng/g compared to 94 ± 11% for Chesapeake Bay.

C.2.5 Sediment Contaminants - Metals

Table C-2 lists minimum, maximum, and median bulk sediment concentrations of metals measured in Chesapeake Bay and Long Island Sound in 1991. Median values for all metals were higher in Long Island Sound than in Chesapeake Bay.

C.2.6 Marine Debris

The incidence of trash collected in trawls is illustrated in Figure C-10. The incidence for both systems was low: trash was found in 15 ± 13% and 1 ± 1% of the area of Chesapeake Bay and Long Island Sound, respectively. The differences between the two systems may be due to the larger number of small estuary and tidal river stations included within the Chesapeake Bay system.

C.3 Habitat Indicators

C.3.1 Light Extinction (Water Clarity)

Water clarity, as defined by human vision, showed definite differences between Chesapeake Bay and Long Island Sound. All of the water of Long Island Sound was classified as "good" relative to clarity based on EMAP sampling (Figure C-11). Approximately 26 ± 11% of the water of Chesapeake Bay was classified as poor or marginal (light extinction coefficient ≥1.387), meaning that a wader could not see his/her toes in waste deep water.

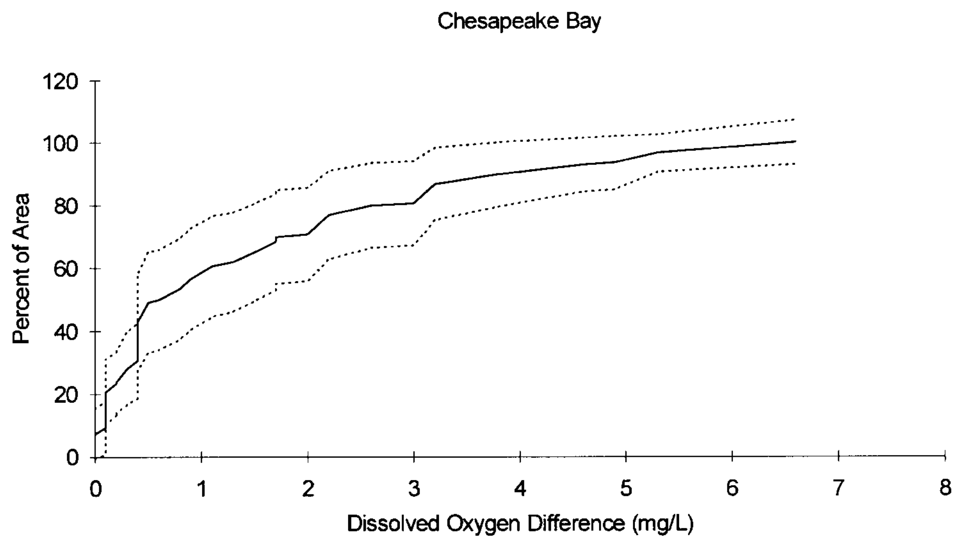
C.3.2 Water Depth

Cumulative distribution functions for water depth in Chesapeake Bay and Long Island Sound are presented in Figure C-12. The Bay is generally much shallower than Long Island Sound. The maximum depths measured in the two systems in 1991 were 22 and 30 m, respectively.

The area shallower than 2 m is underestimated because this is the minimum depth sampled, and, because of the statistical design, unsampleable areas were distributed across the CDF as missing values.

C.3.3 Temperature

The CDFs for bottom water temperature in Chesapeake Bay and Long Island Sound show the Sound to generally contain lower temperature bottom waters than Chesapeake Bay (Figure C-13). This is most likely a function of both water depth and latitude.



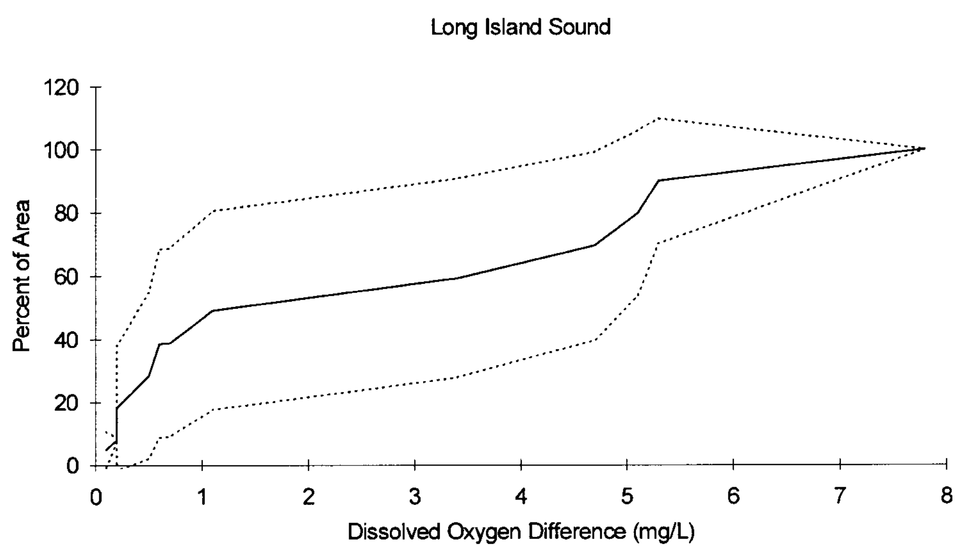
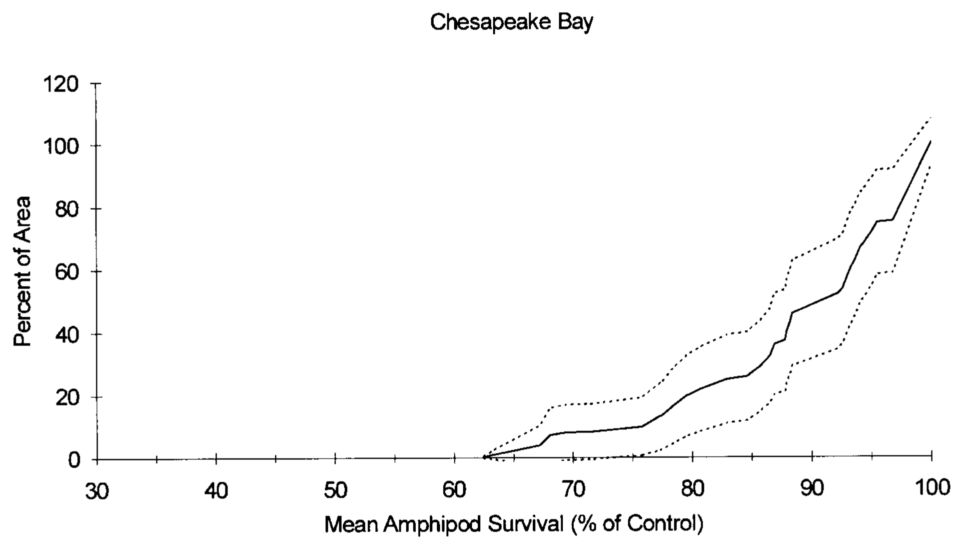


Figure C-7. Cumulative distribution functions of the DO difference between surface and bottom waters as a percent of area of Chesapeake Bay and Long Island Sound, 1991. (Dashed lines are the 95% confidence intervals).



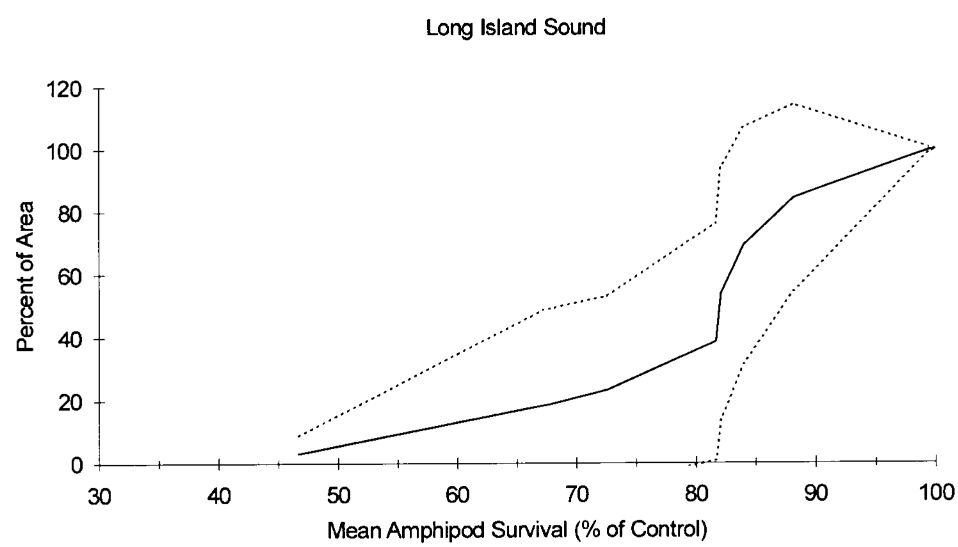


Figure C-8. Cumulative distribution functions of amphipod survival (% of control) in 10-day toxicity tests as a percent of area of Chesapeake Bay and Long Island Sound, 1991. (Dashed lines are the 95% confidence intervals).

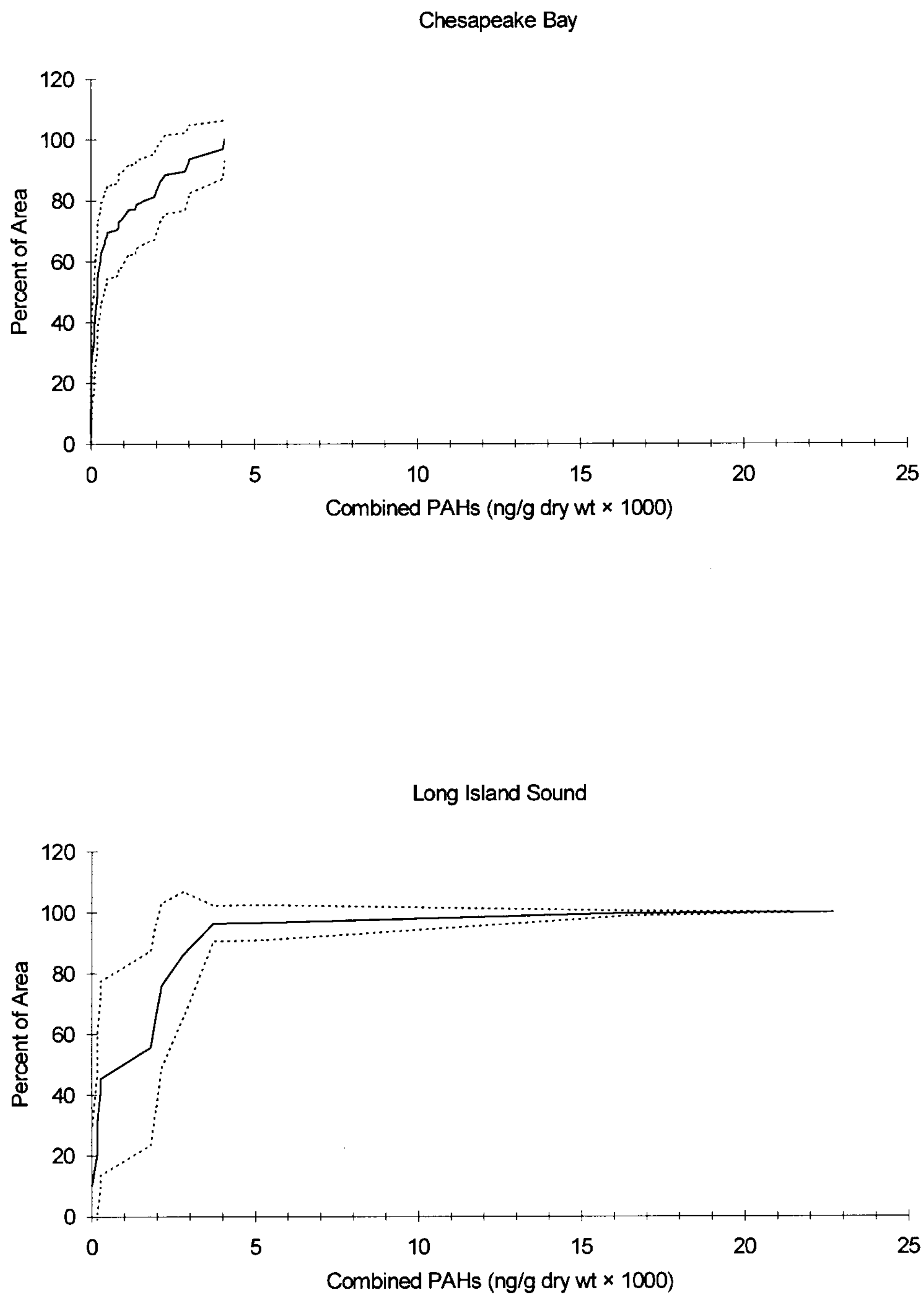


Figure C-9. Cumulative distribution functions of combined PAH concentrations as a percent of area of Chesapeake Bay and Long Island Sound, 1991. (Dashed lines are the 95% confidence intervals).

Table C-2. Range and median metal concentrations in sediments Chesapeake Bay and Long Island Sound, 1991. Concentrations are as µg/g dry weight.

Analyte	MIN	MAX	Median
		Chesapeake Bay	
Major			
Aluminum	1,760	89,300	47,700
Iron	653	54,500	26,300
Manganese	11.6	6,430	364
Trace			
Antimony	ND	1.04	0.262
Arsenic	0.773	21.1	5.34
Cadmium	ND	0.78	0.188
Chromium	1.88	105	42.7
Copper	0.89	47.7	17.5
Lead	ND	274	25.1
Mercury	ND	0.19	0.037
Nickel	ND	50.1	22.8
Selenium	ND	1.76	0.434
Silver	ND	0.52	0.047
Tin	ND	6.16	2.15
Zinc	3.66	266	73.9
		Long Island Sound	
Major			
Aluminum	26,600	65,300	50,500
Iron	8,160	40,300	29,900
Manganese	266	783	593
Trace			
Antimony	0.054	1.40	0.448
Arsenic	1.21	9.43	4.90
Cadmium	0.038	6.58	0.226
Chromium	9.40	174	76.4
Copper	3.10	263	57.5
Lead	11.0	230	45.5
Mercury	0.016	1.96	0.130
Nickel	4.63	46.2	24.0
Selenium	ND	1.41	0.441
Silver	ND	9.69	0.463
Tin	0.362	27.0	5.69
Zinc	22.8	321	145

ND = Not Detected

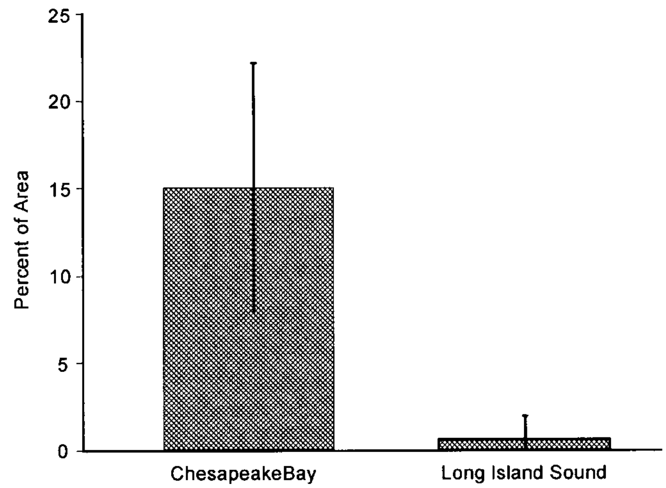


Figure C-10. The incidence of anthropogenic debris in fish trawls as a percent of area of Chesapeake Bay and Long Island Sound, 1991. (Error bars represent the 95% confidence intervals).

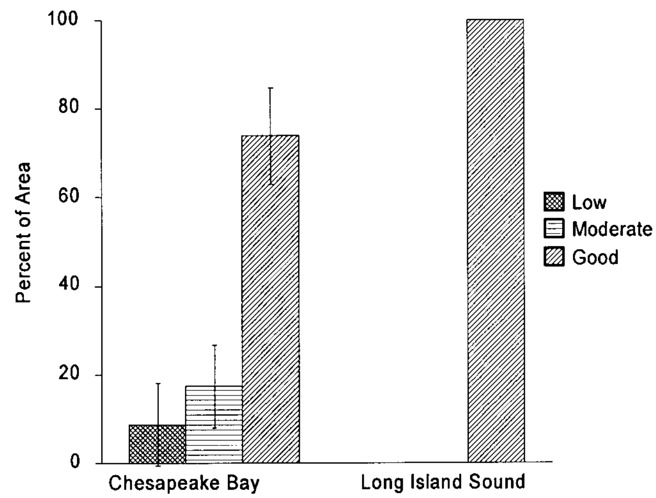
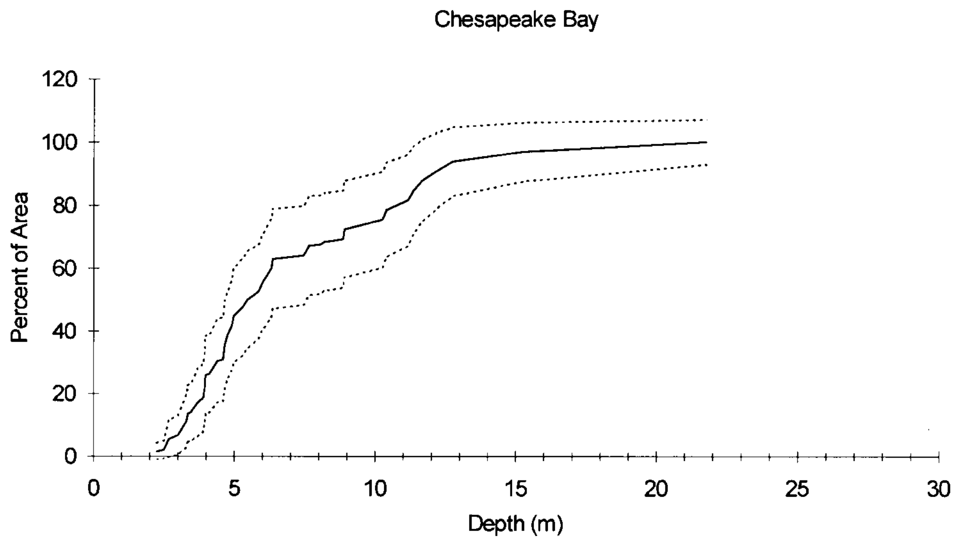


Figure C-11. Percent area of Chesapeake Bay and Long Island Sound with water clarity classified as low, moderate, or good based on light extinction coefficients. (Error bars represent 95% confidence intervals).



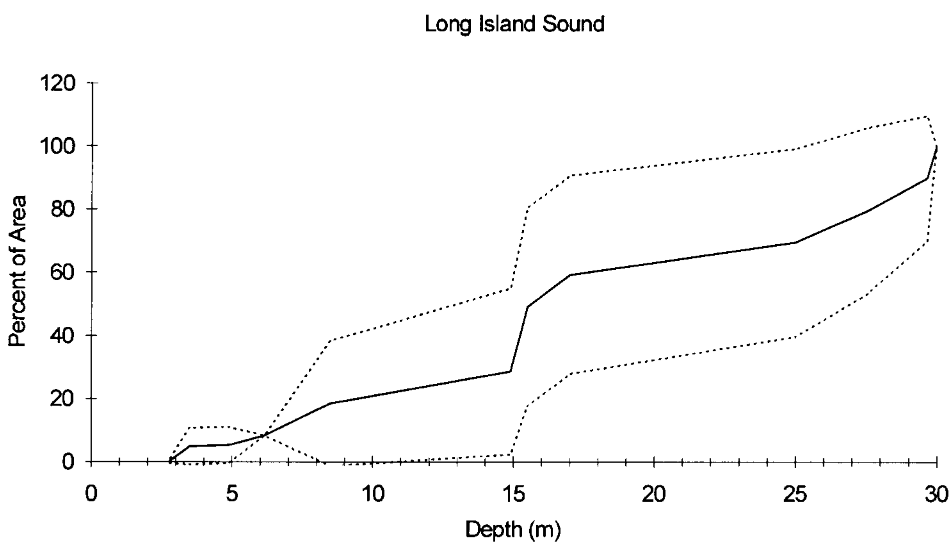
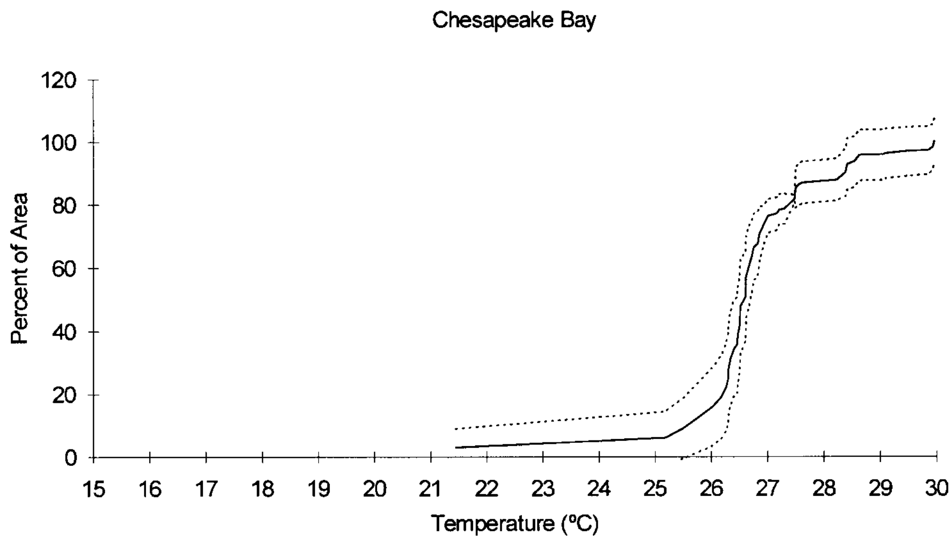


Figure C-12. Cumulative distribution functions of water depth as a percent of area aof Chesapeake Bay and Long Island Sound, 1991. (Dashed lines are the 95% confidence intervals).



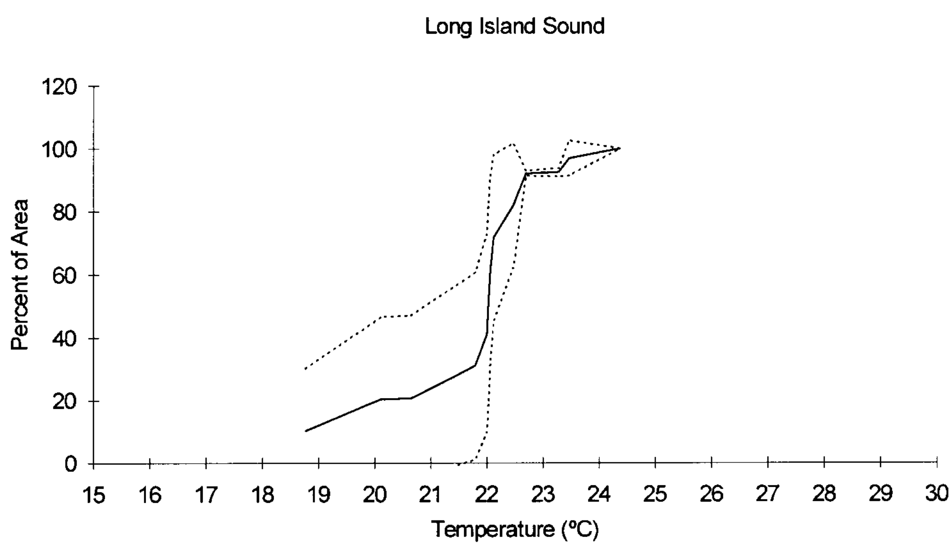


Figure C-13. Cumulative distribution functions of bottom water temperature as a percent of area of Chesapeake Bay and Long Island Sound, 1991. (Dashed lines are the 95% confidence intervals).

C.3.4 Salinity

As described in Section 3, the CDF of salinity in the Virginian Province for 1991 is bimodal, with the major inflections (at 17 and 28 ‰) being accounted for by Chesapeake Bay and Long Island Sound respectively. The combined CDFs for these systems (Figure C-14) illustrate their different salinity patterns.

Long Island Sound contains predominantly polyhaline waters (95 ± 5% of the area sample), with a low salinity tail which is accounted for by one station in the Connecticut River (a small estuary). Chesapeake Bay, because of the inclusion of three major tidal rivers as well as the Susquehanna River, contains a significant area of oligohaline and mesohaline water (64 ± 17% of the area sampled).

C.3.5 Stratification

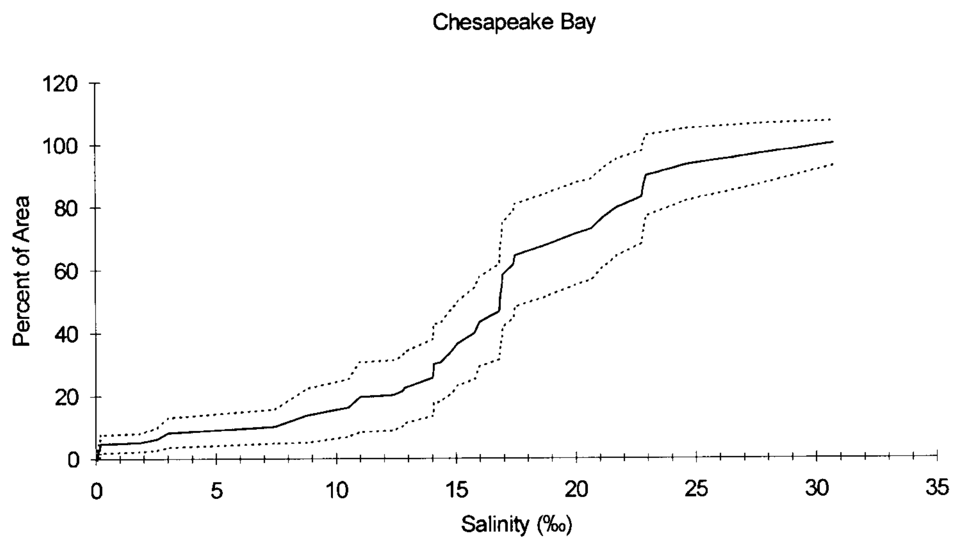
Stratification is shown as CDFs of $\Delta\sigma_t$, which is the σ_t (sigma-t density) difference between surface and bottom waters (Figure C-15).

The greatest stratification in the Province occurred in the lower portion of the Chesapeake Bay. Chesapeake Bay had both the highest area of well-mixed water ($\sigma_t < 1$), and the highest area of significantly stratified water ($\sigma_t > 2$). Most of Long Island Sound fell between σ_t's of 1 and 2.

C.3.6 Percent Silt-Clay Content

The CDFs of silt-clay content for Chesapeake Bay and Long Island Sound are similar with approximately the same percent area of mud and sand in each system (Fig. C-16).

The large area of sandy sediments found in the mouth of the Bay is likely due to sands being carried in from the ocean (Hobbs *et al.*, 1992). In Long Island Sound coarser sediments at the mouth are mainly a result of strong tidal currents transporting away the fine fraction (winnowing), leaving behind the coarser sands and gravel (Akapati, 1974; Gordon, 1980).



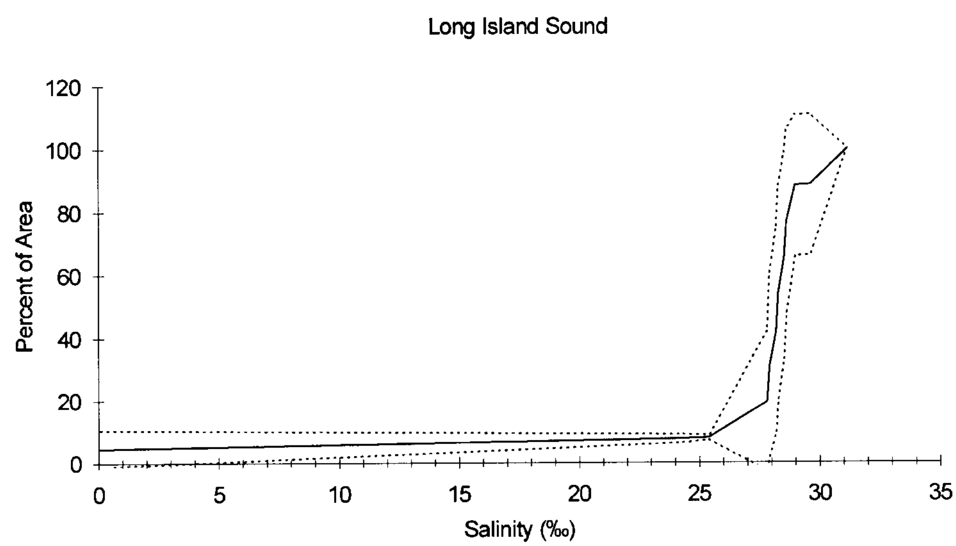
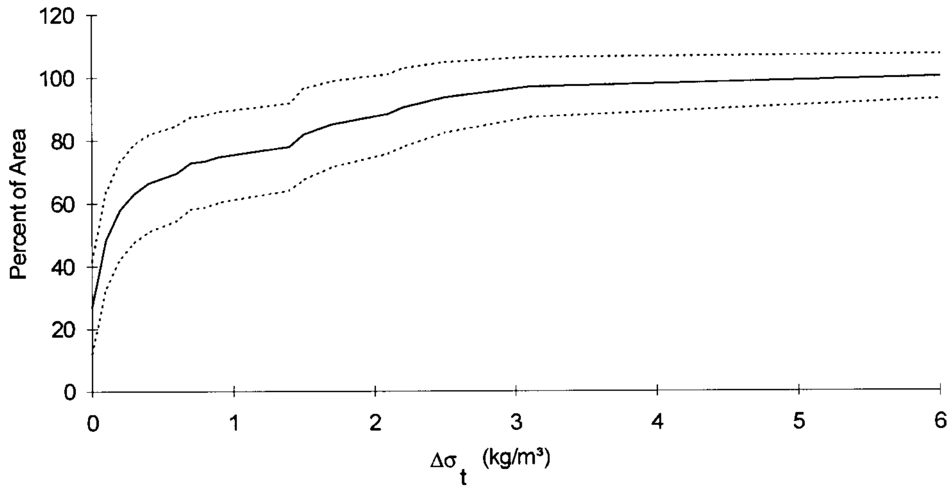


Figure C-14. Cumulative distribution functions of bottom water salinity as a percent of area of Chesapeake Bay and Long Island Sound, 1991. (Dashed lines are the 95% confidence intervals).

Chesapeake Bay



Long Island Sound

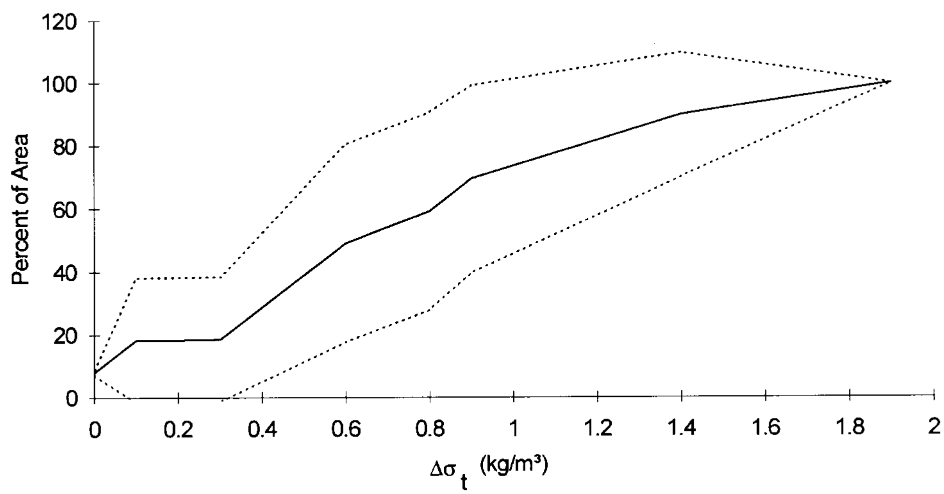
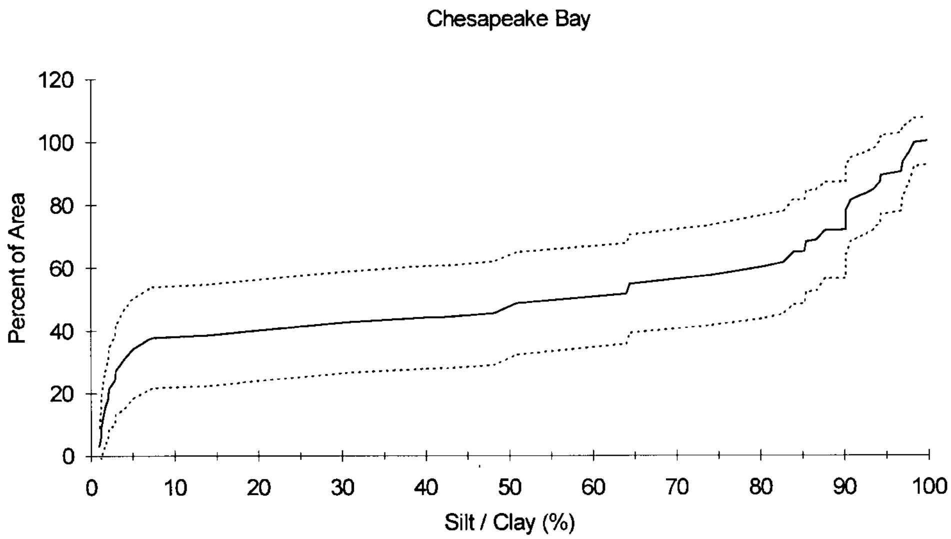


Figure C-15. Cumulative distribution functions of surface to bottom sigma-t density difference as a percent of area of Chesapeake Bay and Long Island Sound, 1991. (Dashed lines are the 95% confidence intervals).



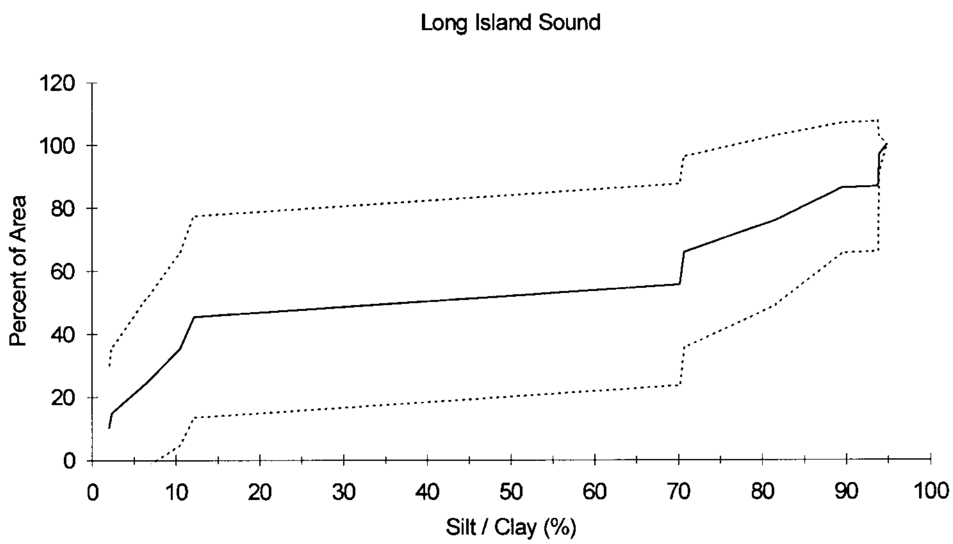


Figure C-16. Cumulative distribution functions of sediment silt/clay content as a percent of area of Chesapeake Bay and Long Island Sound, 1991. (Dashed lines are the 95% confidence interval).

APPENDIX D

LINEAR REGRESSIONS OF INDIVIDUAL METALS AGAINST ALUMINUM USED IN THE DETERMINATION OF METALS ENRICHMENT OF SEDIMENTS OF THE VIRGINIAN PROVINCE

As discussed in Section 3.2.3.7, concentrations of individual metals were normalized against the crustal element aluminum in an attempt to provide a basis for estimating the areal extent of enrichment of these metals in Virginian Province sediments. The method utilized is described in Appendix A (Section A.8.2.3). For each metal, a regression and an upper 95% confidence interval was determined and plotted (Figures D-1 to D-14). Stations with concentrations falling above the upper 95% confidence interval were classified as enriched for that metal. Regression parameters (slope, intercept, and correlation coefficient) are listed in Table D-1.

The results obtained from the regression analyses should be similar to those obtained by other investigators using the same approach. For most of the metals analyzed in the Virginian Province 1991 samples, the results agree exceptionally well with those of Hanson *et al.* (1993) (Table D-1). Correlations between metals and aluminum in 1991 Virginian Province sediments, as measured by the r^2 values obtained, are comparable to those determined from Atlantic and Gulf of Mexico estuarine and coastal sediments collected between 1984 and 1987 by Hanson *et al.* (1993). For eight of the 13 elements analyzed (Hg, Ag, As, Cr, Ni, Sn, Mn and Fe), slopes of the regression lines agree within 50%, while those obtained for other metals (Cu, Cd, Pb, Zn and Se) are 2-3 times higher than found by Hanson *et al.* (1993). In addition, Hanson found statistically significant non-zero intercepts for Ag, Cd, Cr and Pb, indicating significant concentrations of the metals in the silica end-member of the mixing model, or the widespread introduction of metals to all samples within the region samples, *e.g.*, atmospheric deposition of lead. In contrast, none of the intercepts determined from the metal-aluminum regressions in this work are statistically significant.

Agreement between this work and the results of Windom *et al.* (1989) is not as good, but in that study, fewer metals were determined and samples were collected from more limited geographical regions.

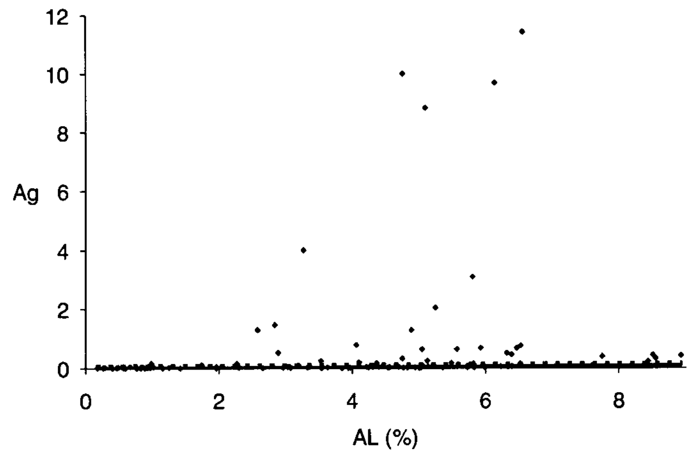


Figure D-1. Linear regression of silver against aluminum (dashed line is the upper 95% confidence interval). Metal concentrations are as µg/g dry weight.

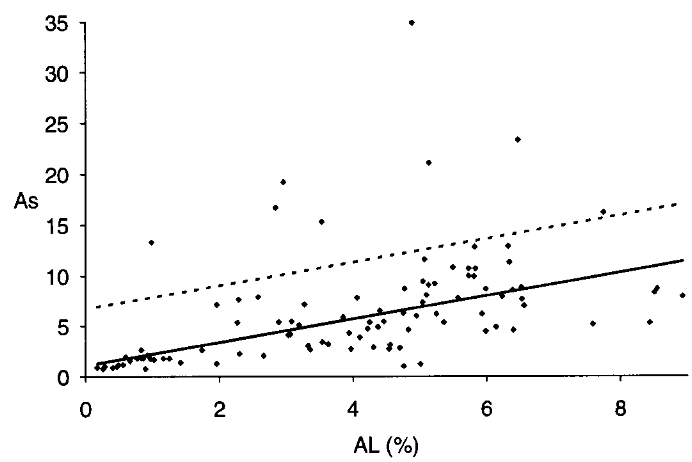


Figure D-2. Linear regression of arsenic against aluminum (dashed line is the upper 95% confidence interval). Metal concentrations are as µg/g dry weight.

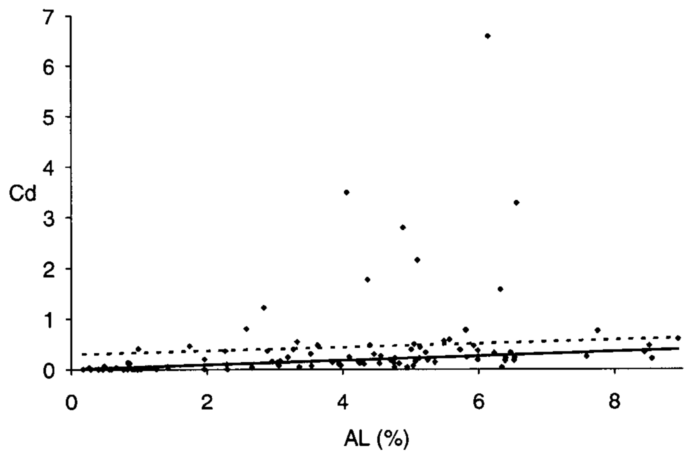


Figure D-3. Linear regression of cadmium against aluminum (dashed line is the upper 95% confidence interval). Metal concentrations are as µg/g dry weight.

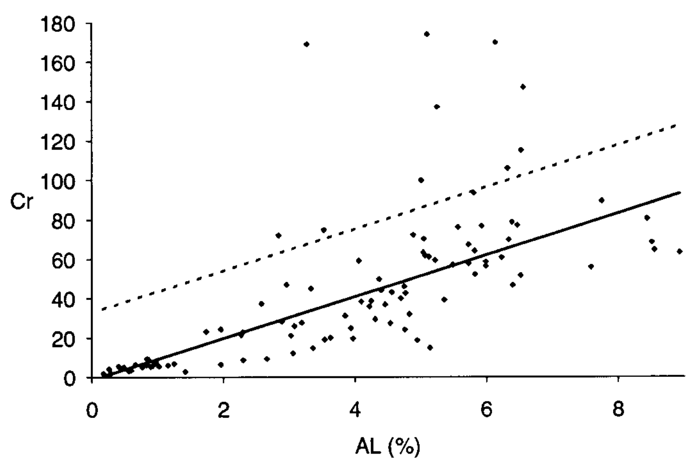


Figure D-4. Linear regression of chromium against aluminum (dashed line is the upper 95% confidence interval). Metal concentrations are as µg/g dry weight.

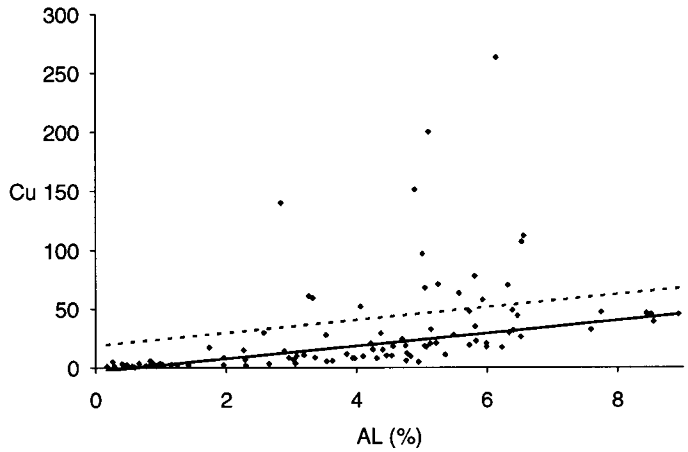


Figure D-5. Linear regression of copper against aluminum (dashed line is the upper 95% confidence interval). Metal concentrations are as µg/g dry weight.

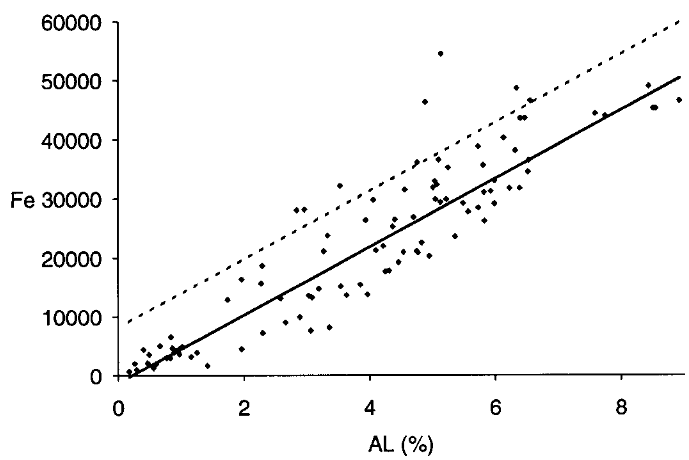


Figure D-6. Linear regression of iron against aluminum (dashed line is the upper 95% confidence interval). Metal concentrations are as µg/g dry weight.

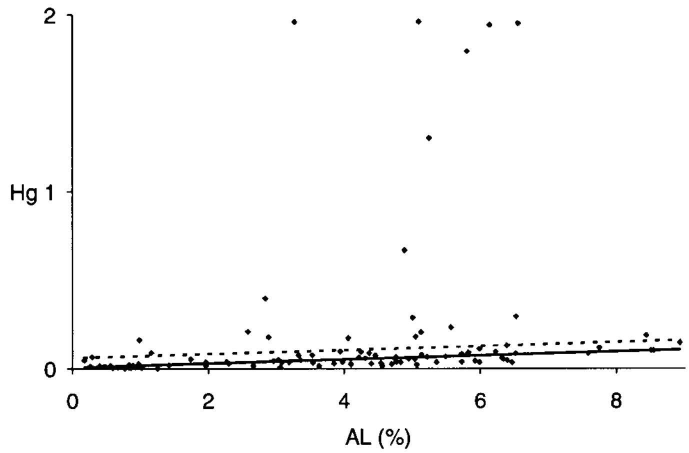


Figure D-7. Linear regression of mercury against aluminum (dashed line is the upper 95% confidence interval). Metal concentrations are as µg/g dry weight.

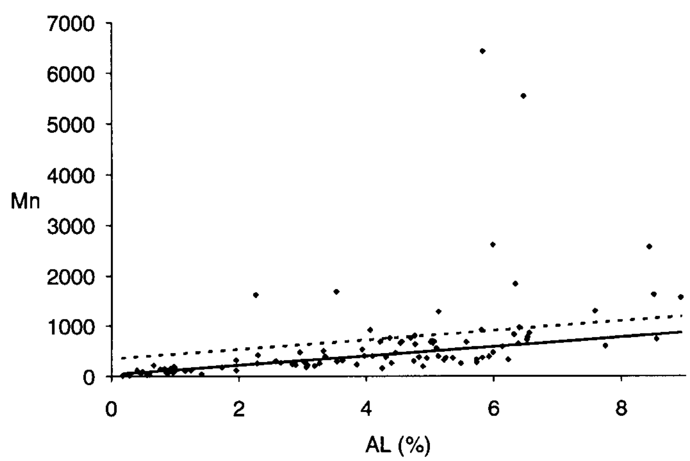


Figure D-8. Linear regression of manganese against aluminum (dashed line is the upper 95% confidence interval). Metal concentrations are as µg/g dry weight.

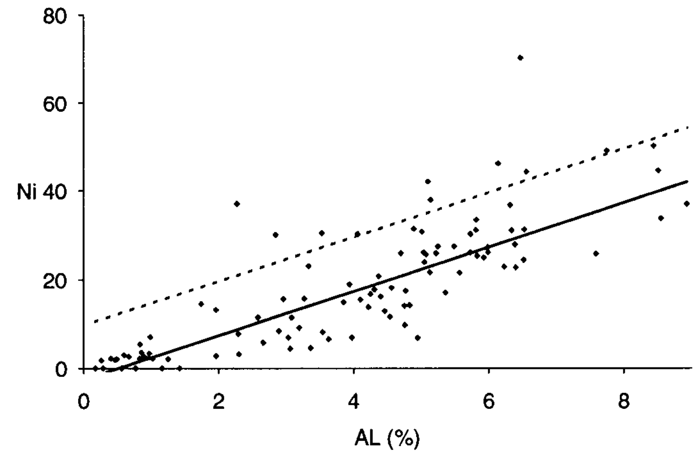


Figure D-9. Linear regression of nickel against aluminum (dashed line is the upper 95% confidence interval). Metal concentrations are as µg/g dry weight.

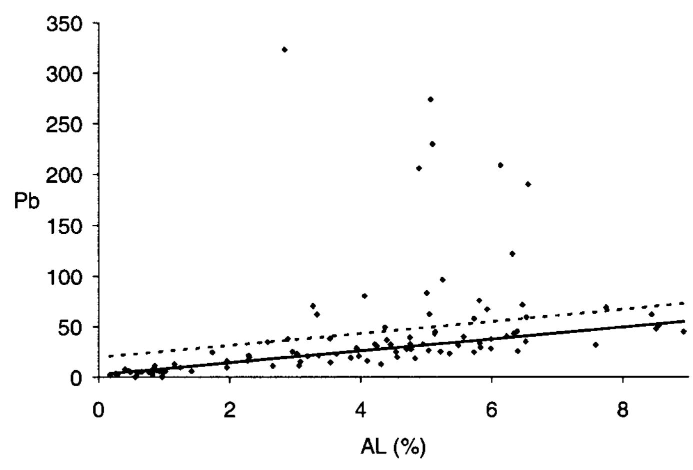


Figure D-10. Linear regression of lead against aluminum (dashed line is the upper 95% confidence interval). Metal concentrations are as µg/g dry weight.

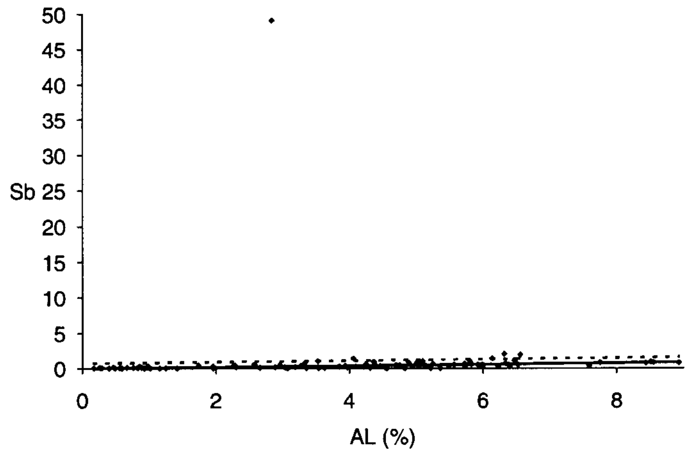


Figure D-11. Linear regression of antimony against aluminum (dashed line is the upper 95% confidence interval). Metal concentrations are as µg/g dry weight.

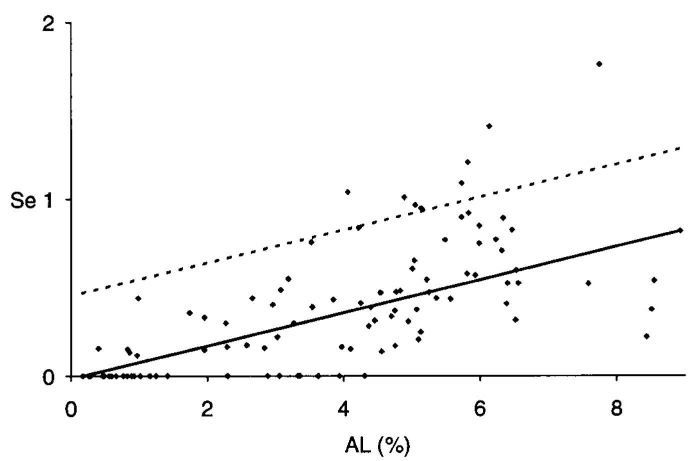


Figure D-12. Linear regression of selenium against aluminum (dashed line is the upper 95% confidence interval). Metal concentrations are as µg/g dry weight.

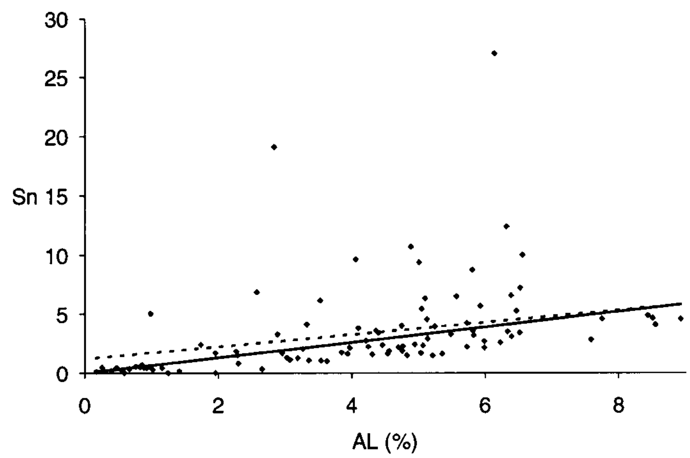


Figure D-13. Linear regression of tin against aluminum (dashed line is the upper 95% confidence interval). Metal concentrations are as µg/g dry weight.

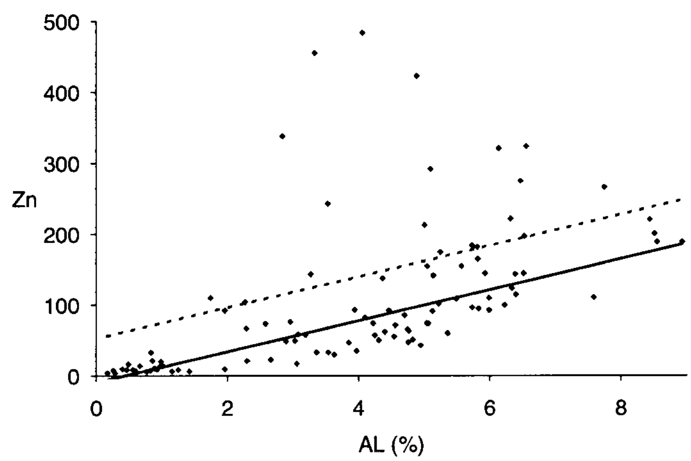


Figure D-14. Linear regression of zinc against aluminum (dashed line is the upper 95% confidence interval). Metal concentrations are as µg/g dry weight.

Table D-1. Comparison of metal-aluminum regression parameters obtained from 1991 Virginian Province sediment data and values reported in the scientific literature (m = slope, b = intercept, r^2 = correlation coefficient).

Element	References	Regression parameters		
		m	b	r^2
Ag	EMAP Va. Prov. - 1991	0.0097	0.0101	0.1939
	Hanson *et al.*, 1993	0.0085	0.0252	0.128
As	EMAP Va. Prov. - 1991	1.15	1.12	0.4706
	Hanson *et al.*, 1993	1.52	0.05	0.606
	Windom *et al.*, 1989 ‡	7.50	-0.70	0.77
Cd	EMAP Va. Prov. - 1991	0.0436	0.0062	0.4527
	Hanson *et al.*, 1993	0.0180	0.0517	0.137
Cr	EMAP Va. Prov. - 1991	10.57	-1.38	0.6661
	Hanson *et al.*, 1993	8.13	9.76	0.635
	Windom *et al.*, 1989 ‡	9.50	4.00	0.81
Cu	EMAP Va. Prov. - 1991	5.44	-3.36	0.5391
	Hanson *et al.*, 1993	1.98	-0.23	0.780
	Windom *et al.*, 1989 *	1.8	-1.4	0.64
	Windom *et al.*, 1989 ‡	2.5	2.2	0.61
Fe	EMAP Va. Prov. - 1991	5788	-1307	0.8879
	Hanson *et al.*, 1993	4950	-612	0.818
	Windom *et al.*, 1989 *	4700	-800	0.91
	Windom *et al.*, 1989 ‡	4800	700	0.88
Hg	EMAP Va. Prov. - 1991	0.0113	0.0057	0.5017
	Hanson *et al.*, 1993	0.0113	-0.0084	0.204
Mn	EMAP Va. Prov. - 1991	93.9	29.3	0.6102
	Hanson *et al.*, 1993	90.1	21.6	0.244
	Windom *et al.*, 1989 *	55	57	0.61
	Windom *et al.*, 1989 ‡	46	27	0.50
Ni	EMAP Va. Prov. - 1991	4.99	-2.62	0.7736
	Hanson *et al.*, 1993	3.31	-1.15	0.633
	Windom *et al.*, 1989 *	4.4	-3.0	0.53
	Windom *et al.*, 1989 ‡	2.9	2.0	0.68
Pb	EMAP Va. Prov. - 1991	5.90	2.39	0.7274
	Hanson *et al.*, 1993	2.99	3.40	0.76
	Windom *et al.*, 1989 *	3.50	1.50	0.62
	Windom *et al.*, 1989 ‡	3.20	2.30	0.69

(continued)

Table D-1 (continued).

Element	References	Regression parameters		
		m	b	r^2
Sb	EMAP Va. Prov. - 1991	0.1008	-0.0011	0.3213
Se	EMAP Va. Prov. - 1991	0.0922	-0.0012	0.4538
	Hanson *et al.*, 1993	0.048	0.002	0.282
Sn	EMAP Va. Prov. - 1991	0.5064	-0.0212	0.7677
	Hanson *et al.*, 1993	0.367	0.292	0.437
Zn	EMAP Va. Prov. - 1991	21.9	-9.60	0.7294
	Hanson *et al.*, 1993	11.7	0.9	0.720
	Windom *et al.*, 1989 *	12	-8	0.70
	Windom *et al.*, 1989 ∓	12	1	0.83

(*) = Data for Georgia and South Carolina coastal sediments
(∓) = Data for Florida coastal sediments